Logical Frameworks for Multiagent Aggregation

Umberto Grandi

ESSLLI, Tübingen, August 2014

Contents

Disclaimer: these are lecture notes for a course given at the *European Summer School on Logic, Language and Information (ESSLLI)* in Tübingen in August 2014. They may be rather hard to read without having attended the course, but they contain a complete list of all definitions and theorems and exercises proposed in class. For questions, solutions, clarifications please contact me by email (umberto.uni@gmail.com).

Chapter 1

Frameworks for Multiagent Aggregation

In the first lecture we go straight into the problem of aggregation: merging individual expressions into a collective view. We see a preview of preference aggregation on 2 alternatives, the formal definition of preference aggregation for any number of alternatives, and the statement of Arrow's Theorem. We look at the definitions of three logical frameworks for aggregation: judgment aggregation, binary aggregation, and binary aggregation with integrity constraints. We conclude with some computational considerations about which framework to choose.

1.1 Condorcet Paradox

We are in the Enlightment period in France: an important problem is the definition of "what do the people want". The *Marquis de Condorcet* discovers a problem in the most straightforward definition, the one using the majority rule.

Consider the following problem:

Alternatives: $\{\triangle, \bigcirc, \square\}$

Individuals: the population of Paris is divided into three equally represented groups: *East, West, South*

Individual expressions: rankings from the most preferred alternative to the least preferred

The problem to solve is: what do the people want? A straightforward possibility (a generalisation of the "head count" used for centuries) is pairwise majority: an alternative is collectively preferred to a second one if a majority of the population prefer the first alternative to the second. But look at the following table:

East	$\triangle < \bigcirc < \square$
West	$\square < \triangle < \bigcirc$
South	$\bigcirc < \square < \triangle$
The People	$\triangle < \bigcirc < \square < \triangle$

The result of the majority rule is a cycle! Any alternative we choose as representing "the will of the people", there will be a majority of the people that prefers another one (i.e., a potential revolution!).

1.2 May's Theorem

Let's try to tackle the problem by first reducing the number of alternatives. It may be that the problem lies in the fact that there are three possible options to choose from (the answer is: yes!).

Consider therefore the following problem:

Alternatives: $\{x, y\}$

Individuals: $\mathcal{N} = \{1, \ldots, n\}$

Individual expressions: $\mathcal{D} = \{x > y, y > x, x \sim y\}$

Where $x > y$ stands for "the individual prefers x to y" and $x \sim y$ for "the individual is indifferent between x and y". Let a *profile* be $\boldsymbol{D} = (D_1, \ldots, D_n)$ with $D_i \in \mathcal{D}$, i.e., the choice of an expression for each of the individuals in $\mathcal{N}$.

Definition 1.1. *An* aggregation procedure *is a function $F : \mathcal{D}^{\mathcal{N}} \to \mathcal{D}$, associating a collective view $F(\boldsymbol{D}) \in \mathcal{D}$ to every profile of individual expressions $\boldsymbol{D}$.*

Some more pieces of notation: let us identify $x > y$ with "1", $y > x$ with "-1", and $x \sim y$ with "0". Let $N(1) = |\{i \in \mathcal{N} \mid D_i = 1\}|$, and similarly for $N(-1)$ and $N(0)$. Here are some examples of aggregation procedures:

- Dictatorships: for all $\boldsymbol{D}$ we have that $F(\boldsymbol{D}) = D_i$
- Constants: for all $\boldsymbol{D}$ we have that $F(\boldsymbol{D}) = 1$
- The majority rule:

$$Maj(\boldsymbol{D}) = \begin{cases} 1 & \text{if } N(1) > N(-1) \\ -1 & \text{if } N(-1) > N(1) \\ 0 & \text{if } N(1) = N(-1) \end{cases}$$

We now ask ourselves the following question: what is a good aggregation procedure? We list below a number of desirable properties that we would like our procedure to satisfy:

Universal domain (hidden): F is defined for all profiles $\boldsymbol{D} \in \mathcal{D}^{\mathcal{N}}$

Anonymity (A): F is a symmetric functions of its arguments, i.e., for all permutations $\sigma : \mathcal{N} \to \mathcal{N}$, we have that $F(D_{\sigma(1)}, \dots, D_{\sigma(n)}) = F(D_1, \dots, D_n)$

Neutrality (N): F is an odd function, i.e., $F(-D_1, \dots, -D_n) = -F(D_1, \dots, D_n)$

Positive Responsiveness (PR): If $F(D_1, \dots, D_n) = 1$ or $= 0$, and $D'_i = D_i$ for all $i \neq \bar{i}$, and $D'_{\bar{i}} > D_{\bar{i}}$, then $F(D'_1, \dots, D'_n) = 1$.

Is there any "good" procedure? One that satisfies all these conditions? Yes, there is actually only one:

Theorem 1.1 (May, 1952)**.** *An aggregation procedure satisfies A, N and PR, if and only if it is the majority rule.*

Proof. $\Leftarrow$) It is easy to see that Maj satisfies the three desired axioms: A is satisfied since Maj is defined in terms of $N(1)$ and $N(-1)$ only; N is satisfied since

$$Maj(-\boldsymbol{D}) = 1 \Leftrightarrow N(-1) > N(1) \Leftrightarrow Maj(\boldsymbol{D}) = -1;$$

and, finally, PR is satisfied since when $Maj(\boldsymbol{D}) = 0$ then $N(1) = N(-1)$, hence in $\boldsymbol{D'}$ we have that $N(1) > N(-1)$ and so $Maj(\boldsymbol{D'}) = 1$.

$\Rightarrow$) The interesting part is showing that the majority is the only rule satisfying these axioms. Let F satisfy A, N and PR. By A, we know that there exists a function $h : \{1, \dots, n\}^2 \to \mathcal{D}$ such that $Maj(\boldsymbol{D}) = h(N(1), N(-1))$. We first cover the last line of the definition of Maj:

$$\begin{aligned} N(-1) = N(1) \Rightarrow \\ F(\boldsymbol{D}) = h(N(1), N(-1)) = h(N(-1), N(1)) = F(-\boldsymbol{D}) \\ \Rightarrow F(\boldsymbol{D}) = 0 \end{aligned}$$

Let us now consider the case in which:

$$N(1) = N(-1) + 1 \Rightarrow F(\boldsymbol{D}) = 1$$

using PR starting from a profile where $N(1) = N(-1)$. Now an inductive use of PR shows that, when $0 \leqslant m \leqslant n - N(1)$:

$$N(1) = N(-1) + m \quad \Rightarrow \quad F(\boldsymbol{D}) = 1$$

Hence if $N(1) > N(-1)$ we obtain that $F(\boldsymbol{D}) = 1$, taking care of the first case in the definition of the majority rule. If is then sufficient to use the neutrality property N and conclude that:

$$N(-1) > N(1) \Rightarrow F(-\boldsymbol{D}) = 1 \Rightarrow F(\boldsymbol{D}) = -1$$

thus concluding the proof. $\square$

Exercise 1.1. *Show that the three properties A, N and PR are independent, i.e., that there exists aggregation procedures F that satisfy only one of the three properties but not the other two.*

References: the original paper by May (1952).

1.3 Preference Aggregation

May's Theorem showed that if we want to aggregate the preferences of a collectivity over two alternatives, then the majority rule is the only aggregation procedure that satisfies a number of desirable properties. The Condorcet paradox showed instead that with three (or more?) alternatives this rule may result in unexpected outcomes. We now formalise the setting of preference aggregation for any number of alternatives, and give the statement of a famous theorem showing that situations like the Condorcet paradox cannot be easily avoided.

The setting is the following:

Alternatives: $\mathcal{X}$

Individuals: $\mathcal{N} = \{1, \dots, n\}$

Individual expressions: the set of all linear orders $\mathcal{L}(\mathcal{X})$ (i.e., anti-simmetric, transitive and complete binary relations over $\mathcal{X}$)

A profile of individual preferences is $\boldsymbol{P} = (P_1, \dots, P_n)$ where each $P_i \in \mathcal{L}(\mathcal{X})$ is a linear order. Clearly we could consider different ways of expressing preferences, keeping the assumption that they will be binary relations: weak orders (reflexive, transitive and complete relations on $\mathcal{X}$), partial orders (reflexive and transitive relations), dichotomous preferences (subsets $P \subset \mathcal{X}$ of "approved" alternatives). In this notes we will restrict to linear orders (but keep this in mind in Lecture 2 and 4).

Definition 1.2. *An* aggregation procedure *is a function $F : \mathcal{L}(\mathcal{X})^{\mathcal{N}} \to \mathcal{L}(\mathcal{X})$, associating a collective preference $F(\boldsymbol{P}) \in \mathcal{L}(\mathcal{X})$ to every profile of individual expressions $\boldsymbol{P}$.*

We write xP_iy for "individual i prefers alternative x to alternative y, similarly for the collective preference $F(\boldsymbol{P})$. Here are some examples of aggregation procedures:

- Dictatorships: for all $\boldsymbol{P}$ we have that $F(\boldsymbol{P}) = P_i$
- Constants: for all $\boldsymbol{P}$ we have that $F(\boldsymbol{P}) =<$ for a given linear order $<$
- The majority rule (is this a proper aggregation procedure?): for all $x, y \in \mathcal{X}$ we define $xMaj(\boldsymbol{P})y$ iff $|\{i \in \mathcal{N} \mid xP_iy\}| > \lceil \frac{n+1}{2} \rceil$

Once more, let us think of properties that would make for a "good" aggregation procedure. Similar properties as those we expressed in the previous section can be written, but we will focus on a set of weaker properties (see Exercise 1.2).

Universal domain: (hidden) F is defined for every profile $\boldsymbol{P}$

Collective rationality: (hidden) the outcome of F is a linear order, i.e., $F(\boldsymbol{P}) \in \mathcal{L}(\mathcal{X})$ for every profile $\boldsymbol{P}$. Does the majority rule satisfy this requirement?

Unanimity (U): if for all $i \in \mathcal{N}$ we have that xP_iy then also $xF(\boldsymbol{P})y$ (aka. weak-Pareto property)

Independence of Irrelevant Alternatives (IIA): for all profiles $\boldsymbol{P}$ and $\boldsymbol{P'}$, if for all $i \in \mathcal{N}$ we have that $xP_iy \Leftrightarrow xP'_iy$ then also $xF(\boldsymbol{P})y \Leftrightarrow xF(\boldsymbol{P'})y$

Non-dictatorship (NDIC): F is not a dictatorship, i.e., there is no $i \in \mathcal{N}$ such that for all profiles $\boldsymbol{P}$ we have that $F(\boldsymbol{P}) = P_i$.

We have seen that the majority rule is not a candidate for a good procedure, since by the Condorcet paradox it does not even output a linear order. The surprising result is that there is no rule that satisfies the properties above!

Theorem 1.2 (Arrow,1951)**.** *Let $|\mathcal{X}| \geqslant 3$ and $|\mathcal{N}| \geqslant 2$, there is no aggregation procedure that satisfies U, IIA and NDIC. (Usually written as "every unanimous and independent aggregation procedure is dictatorial").*

Proof. Stay tuned until Lesson 4. □

Instead of proving the theorem let us discuss about possible logical formalisations of this theorem. A formalisation in first-order logic has been given by Grandi and Endriss (2013b), showing that Arrow's statement is equivalent to a given first-order theory not having any finite model. Formalisations in modal logic have been presented by Troquard et al. (2011) and Ågotnes et al. (2009). An alternative approach is the one followed by Tang and Lin (2009), where Arrow's statement is proven using a SAT solver, after having been reduced to the base case of 3 alternatives and 2 individuals with an inductive lemma.

Exercise 1.2. *State the A, N and PR axioms for the general setting of preference aggregation. Show that for the case of 2 alternatives those axioms are stronger than Arrow's axioms U, IIA and NDIC, i.e. there are aggregation procedures that satisfies the latter ones but not the former. Is this true also for more than 3 alternatives?*

References: the beautiful textbook by Moulin (1988); other introductory books on the topic have been written by Gaertner (2006) and by Taylor (2005); the original publication by Arrow (1963) is still actual and is a pleasure to read; another classic is Sen (1970), where every mathematical chapter is preceded by an explanatory one.

1.4 Judgment Aggregation

The world of individual expressions is not only made of preferences. People express more complex notions, and we will now make a step forward by generalising from preferences to judgments about logically related propositions. We define in the following sections three frameworks for a more general study of aggregation, more or less following the order of historical appearance.

Let $\mathcal{L}$ be a set of propositional formulas built from a finite set of propositional variables using the usual connectives $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$, and the constants $\top$ and $\bot$. For every formula α, define $\sim\alpha$ to be the *complement* of α, i.e., $\sim\alpha := \neg\alpha$ if α is not negated, and $\sim\alpha := \beta$ if $\alpha = \neg\beta$ for some formula β. We say that a set Φ is *closed under complementation* if it is the case that $\sim\alpha \in \Phi$ whenever $\alpha \in \Phi$. The setting of jugment aggregation is the following:

Agenda: $\Phi \subseteq \mathcal{L}_{PS}$, finite and nonempty, closed under complementation, and not containing any double negation.

Individuals: $\mathcal{N} = \{1, \ldots, n\}$

Individual expressions: judgment sets $J \subseteq \Phi$

Rationality: for all $\varphi \in \Phi$ we have that $\varphi \in J$ or $\sim \varphi \in J$ (complete), and such that $\bigwedge_{\varphi \in J} \varphi$ has a model (consistent)

Example 1.1. *Let $\Phi = \{p, \neg p, q, \neg q, p \wedge q, \neg(p \wedge q)\}$. A complete and consistent judgment set is $J = \{p, \neg q, \neg(p \wedge q)\}$.*

Let $\mathcal{J}(\Phi)$ denote the set of all complete and consistent subsets of Φ. A profile of judgments is $\boldsymbol{J} = (J_1, \ldots, J_n)$ such that $J_i \in \mathcal{J}(\Phi)$.

Definition 1.3. *Given a finite agenda Φ and a finite set of individuals $\mathcal{N}$, an* aggregation procedure *for Φ and $\mathcal{N}$ is a function $F : \mathcal{J}(\Phi)^{\mathcal{N}} \rightarrow 2^{\Phi}$, associating a collective judgment set $F(\boldsymbol{J}) \subseteq \Phi$ to every profile of individual expressions $\boldsymbol{J}$.*

We can also give properties of a "good aggregation procedure":

Universal domain: (hidden) as usual...

Collective rationality: (not hidden!) $F(\boldsymbol{J})$ is consistent and complete

Other properties...

Exercise 1.3. *Write properties such as A, N, PR, IIA, U, NDIC for the setting of judgment aggregation. Write the definition of the majority rule and test which properties are satisfied (is your definition of the majority rule an aggregation procedure?).*

References: a very recent textbook on judgment aggregation has been written by Grossi and Pigozzi (2014) based on an ESSLLI course; a shorter introduction has been written by List and Puppe (2009); and, of course, the original paper by List and Pettit (2002).

1.5 Binary Aggregation

The setting of judgment aggregation looks complex. Let's try to simplify it in the following way:

Issues: a set of binary issues $\mathcal{I} = \{1, \ldots, m\}$

Individuals: $\mathcal{N} = \{1, \ldots, n\}$

Individual expressions: binary ballots $B \in \mathcal{D} = \{0, 1\}^{\mathcal{I}}$, expressing yes/no opinions on the issues

Rationality: a set $X \subseteq \{0, 1\}^{\mathcal{I}}$ of admissible ballots

A profile or *rational* ballots is $\boldsymbol{B} = (B_1, \ldots, B_n)$ such that $B_i \in X$. There is a new entry of rationality, since issues are not always independent and thus not all views are admissible! Again we can define an aggregation procedure (and again Exercise 1.3 applies).

References: initially introduced by Wilson (1975), it was subsequently expanded by Dokow and Holzman (2010, 2009)

1.6 Binary Aggregation with Integrity Constraints

The idea of listing all admissible views sounds weird to a computer scientist. We can specify admissible ballots in a compact way by using a constraint:

Issues: a set of binary issues $\mathcal{I} = \{1, \ldots, m\}$

Individuals: $\mathcal{N} = \{1, \ldots, n\}$

Individual expression: binary ballots $B \in \mathcal{D} = \{0, 1\}^{\mathcal{I}}$, expressing yes/no opinions on the issues

Rationality: a formula $\mathrm{IC} \in \mathcal{L}_{PS}$, where $\mathcal{L}_{PS}$ is the propositional language constructed from atoms $PS = \{p_1, \ldots, p_m\}$

A binary ballot B is admissible if $B \models \mathrm{IC}$, and a profile is $\boldsymbol{B} = (B_1, \ldots, B_n)$ with $B_i \models \mathrm{IC}$ An aggregation procedure can be defined as follows:

Definition 1.4. *An* aggregation procedure *is a function $F : \mathcal{D}^{\mathcal{N}} \to \mathcal{D}$, associating a collective ballot $F(\boldsymbol{B}) \in \mathcal{D}$ to every profile of individual expressions $\boldsymbol{B}$.*

Again, desirable properties can be defined. A very interesting fact is that we can express the property of *collective rationality* depending on the IC:

Definition 1.5. *Given an integrity constraint* $\mathrm{IC} \in \mathcal{L}_{PS}$*, an aggregation procedure $F : \mathcal{D}^{\mathcal{N}} \to \mathcal{D}$ is called* collectively rational *(CR) with respect to* IC*, if for all rational profiles $\boldsymbol{B} \in \mathrm{Mod(IC)}^{\mathcal{N}}$ we have that $F(\boldsymbol{B}) \in \mathrm{Mod(IC)}$.*

References: the simplest introduction is Grandi and Endriss (2011).

1.7 Which framework is (computationally) best? (Extra material)

The last three frameworks, judgment aggregation (JA), binary aggregation (BA), and binary aggregation with integrity constraints (BAwithIC), seems general enough to model a variety of individual expressions. Which one should be pick if we have a computationally oriented mind?

First, we should formally prove that their expressive power is the same. We need therefore a common semantics on which the three frameworks can be interpreted. We choose subsets of feasible binary evaluations as our *common semantics*, i.e. the objects expressed by the four frameworks:

Definition 1.6. *The common semantics for judgments is* $\mathcal{L}_0 = \{X \mid X \subseteq \{0,1\}^m, m \in \mathbb{N}\}$

Given the combinatorial explosion (in terms of m) of $\mathcal{L}_0$, the four frameworks introduced previously can be used to compactly represent elements of $\mathcal{L}_0$.

Definition 1.7. *A* language for judgments $(\mathcal{L}, \tau)$ *is given by a set of objects* $\mathcal{L}$ *and an interpretation function* $\tau : \mathcal{L} \to \mathcal{L}_0$*.*

Exercise 1.4. *Write the three frameworks introduced before as in Definition 1.6, defining* $\mathcal{L}_{JA}, \mathcal{L}_{BA}$ *and* $\mathcal{L}_{IC}$ *as well as* τ_{JA}, τ_{BA} *and* τ_{IC}*. Hint: for BA there is not much to do.*

Now we can finally state (and the reader is encouraged to prove it, at least for BAwithIC and BA) the equivalence in terms of expressivity:

Proposition 1.1. *The four frameworks are all fully expressive.*

The first parameter we use to judge whether a framework is computationally the best is *succintness*, i.e., how compact is the representation of the same object in the various frameworks:

Definition 1.8. $\mathcal{L}$ *is* at least as succinct as $\mathcal{L}'$*, and we write* $\mathcal{L} \preceq \mathcal{L}'$ *if there exists a function* $f : \mathcal{L}' \to \mathcal{L}$ *and a polynomial* p *such that:*

- $f(X) \sim X$ *for all* $X \in \mathcal{L}'$*;*
- $\text{size}(f(X)) \leqslant p(\text{size}(X))$ *for all* $X \in \mathcal{L}'$*.*

$\mathcal{L}$ *is* strictly more succinct *than* $\mathcal{L}'$ *if* $\mathcal{L} \preceq \mathcal{L}'$ *but* $\mathcal{L}' \not\preceq \mathcal{L}$*.* $\mathcal{L}$ *is* equally succinct *as* $\mathcal{L}'$ *if* $\mathcal{L} \preceq \mathcal{L}'$ *and* $\mathcal{L}' \preceq \mathcal{L}$*.*

Exercise 1.5. *Show that* $\mathcal{L}_{IC}$ *is strictly more succinct than* $\mathcal{L}_{BA}$*. (easy)*

We can therefore discard binary aggregation, since we have a strictly more compact way to express the same problems.

Exercise 1.6. *Show that* $\mathcal{L}_{IC}$ *is equally succinct as* $\mathcal{L}_{JA}$*, or that it is incomparable. (Open problem,* ***prize:*** *one beer for each* $\preceq$ *shown, 4 beers for incomparability)*

The second parameter we use to judge these frameworks is the computational complexity of some basic tasks that needs to be solved when using them. From the point of view of an agent, a decision problem framed in BAwithIC 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 with IC deciding whether $B_i \models \mathrm{IC}$ 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.

Chapter 2

Paradoxes of Aggregation

Most work in Social Choice Theory started with the observation of paradoxical situations. It is not a coincidence that these lecture notes also started from the Condorcet paradox. 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 second lecture 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 with integrity constraints with a uniform formulation. Such a uniform representation of the most important paradoxes in Social Choice Theory enables us to make a crucial observation and show an important result concerning the syntactic structure of paradoxical integrity constraints for the majority rule: they all feature a disjunction of literals of size at least 3.

2.1 Binary Aggregation with Integrity Constraints

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. Before getting to the paradoxes, we recall the basic definitions of the framework of binary aggregation with integrity constraints, and give a general definition of paradox.

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* $\boldsymbol{B} = (B_1, \dots, B_n)$. 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 $\boldsymbol{B} = (B_1, \dots, B_n)$.

Definition 2.1. *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.

If $\mathcal{I}$ is a set of m issues, let $PS = \{p_1, \dots, p_m\}$ 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 $\mathrm{Mod}(\varphi)$ be the set of assignments that satisfy φ.

Definition 2.2. *An* integrity constraint *is any formula* $\mathrm{IC} \in \mathcal{L}_{PS}$.

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 $p_1, \dots, p_m$, and we call B a *rational ballot* if it satisfies the integrity constraint IC, i.e., if B is an element of $\mathrm{Mod}(\mathrm{IC})$. A rational profile is an element of $\mathrm{Mod}(\mathrm{IC})^{\mathcal{N}}$. 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. *(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 $p_U \wedge p_S \rightarrow p_C$.

Example 2.2. *(Voting for candidates) A winning candidate has to be chosen from a set $C = \{1, \ldots, m\}$ 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:*

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

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.

Exercise 2.1. *(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. Write the situation in binary aggregation terms, and devise the right set of integrity constraints.*

2.1.4 A general definition of paradox

Consider the situation introduced in Example 2.1: There are three issues at stake, and the integrity constraint is represented by the formula 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 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.

Definition 2.3. *A* paradox *is a triple* $(F, \boldsymbol{B}, \mathrm{IC})$*, where* $F : \mathcal{D}^{\mathcal{N}} \to \mathcal{D}$ *is an aggregation procedure,* $\boldsymbol{B}$ *is a profile in* $\mathcal{D}^{\mathcal{N}}$*,* IC *is an integrity constraint in* $\mathcal{L}_{PS}$*, and* $B_i \in \mathrm{Mod(IC)}$ *for all* $i \in \mathcal{N}$ *but* $F(\boldsymbol{B}) \notin \mathrm{Mod(IC)}$*.*

We are now ready to explore the generality of Definition 2.3 by showing that classical paradoxes introduced in several frameworks for multiagent aggregation are instances of this definition.

2.2 The Condorcet Paradox revisited

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). We have already seen a version of the paradox in the first lecture, but let us state it again here:

> **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 2.2.

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

Table 2.2: 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 2.3).

2.2.1 Preference Aggregation

We recall from Lecture 1 the basic definitions of preference aggregation. The framework of *preference aggregation* 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.

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

Definition 2.4. *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 2.2, 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.

2.2.2 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).

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). 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 $\mathcal{X}$.

Exercise 2.2. *Devise the right integrity constraint for weak orders* $\mathrm{IC}_{\leqslant}$ *and partial orders* IC_p.

2.2.3 The Condorcet Paradox in Binary Aggregation

The translation presented in the previous section enables us to express the Condorcet paradox in terms of Definition 2.3. 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 2.3, 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 2.3 satisfies $\mathrm{IC}_<$,

	$\triangle\bigcirc$	$\bigcirc\square$	$\triangle\square$
Agent 1	1	1	1
Agent 2	1	0	0
Agent 3	0	1	0
Maj	1	1	0

Table 2.3: Condorcet paradox in binary aggregation.

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 2.3.

[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.

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.

References: The Condorcet paradox (and preference aggregation more in general) has been interpreted into judgment aggregation by Dietrich and List (2007b) by using a set of fist-order predicates rather than propositional atoms. An interpretation similar to ours has been given by Dokow and Holzman (2010). There are also several publications comparing the strength of judgment aggregation with respect to preference aggregation, among others the work of Porello (2010) and that of Grossi (2009, 2010).

2.3 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*, which we have briefly discussed in the first lecture. 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 2.4.

	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 2.4: 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.

We now show that the discursive dilemma is also an instance of our general definition of paradox.

2.3.1 Judgment Aggregation

Once more, recall from the first lecture the basic definitions of 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. 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 2.5. *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.

2.3.2 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. 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 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 $\mathcal{D}_\Phi^{\mathcal{N}}$.

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).

2.3.3 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 2.5.

	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 2.5: 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, $(Maj, \boldsymbol{B}, \mathrm{IC}_\Phi)$ constitutes a paradox by Definition 2.3.

References: Another interesting interpretation of the doctrinal paradox is the geometrical one given by (Eckert and Klamler, 2009) using the theoretical setting developed by Saari (2008).

2.4 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 2.6.

	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 2.6: 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). 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).

2.4.1 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).[2] A binary ballot over these issues represents the 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 2.7.

	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 2.7: The Ostrogorski paradox in binary aggregation.

Each individual in Table 2.7 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 $(Maj, \boldsymbol{B}, \mathrm{IC}_O)$ constitutes a paradox by Definition 2.3.

Exercise 2.3. *Devise a stronger version of the Ostrogorski paradox, in which the winning party disagrees with a majority of the individuals on* all *issues.*

[2] 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.

2.5 The Majority Rule: Characterisation of Paradoxes

We can now make a crucial observation concerning the syntactic structure of the integrity constraints that formalise the paradoxes we have presented so far. First, for the case of the Condorcet paradox, we observe that the formula encoding the transitivity of a preference relation is the implication $p_{ab} \wedge p_{bc} \to p_{ac}$. 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. Second, the formula which appears in the translation of the discursive dilemma is also equivalent to a clause of size 3, namely $\neg p_\alpha \vee \neg p_\beta \vee \neg p_{\neg(\alpha\wedge\beta)}$. Third, the formula which formalises the majoritarian constraint underlying the Ostrogorski paradox, is equivalent to the following conjunction of clauses of size 3:

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

Thus, we observe that the **integrity constraints formalising the most classical paradoxes in aggregation theory all feature a clause (i.e., a disjunction) of size at least 3.**[3] In this section we characterise the class of integrity constraints that are lifted by the majority rule as those formulas that can be expressed as a conjunction of clauses (i.e., disjunctions) of maximal size 2.

Let us first provide a formal definition of the majority rule.

Definition 2.6. *Let $N_j^{\boldsymbol{B}}$ be the set of individuals that accept issue j in profile $\boldsymbol{B}$. In case the number of individuals is odd, the* majority rule *(*Maj*) has a unique definition by accepting issue j if and only if $|N_j^{\boldsymbol{B}}| \geqslant \frac{n+1}{2}$.*

For the remainder of this section we make the assumption that the number of individuals is odd. Recall from the first lecture the definition of collective rationality:

Definition 2.7. *F is* collectively rational *(CR) wrt.* IC *if it does not generate paradoxes with* IC*, i.e., if $F(\boldsymbol{B}) \models \mathrm{IC}$ whenever for all $i \in N$ we have that $B_i \models \mathrm{IC}$.*

We now show in the following theorem that a necessary and sufficient condition for the majority rule not to generate a paradox with respect to a given integrity constraint IC is that IC be equivalent to a conjunction of clauses of maximal size 2. The kernel of the proof goes on the following lines: first, we show that clauses of size one, i.e., literals, and of size 2 never create a paradox with the majority rule; second, we show that it is always possible to generate a paradox with a clause of size 3 or bigger.

[3] This observation is strongly related to a result 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.

Theorem 2.1. *The majority rule Maj is CR with respect to* IC *if and only if* IC *is equivalent to a conjunction of clauses of maximal size 2.*

Proof. ($\Leftarrow$) Let IC be equivalent to a conjunction of clauses of maximal size 2, which we indicate as $\psi = \bigwedge_k D_k$. We want to show that *Maj* is CR wrt. IC. We first make the following two observations. First, since two equivalent formulas define the same set of rational ballots, showing that *Maj* is CR wrt. IC is equivalent to showing that *Maj* is CR wrt. ψ. Second, if the majority rule is collectively rational wrt. two formulas φ_1 and φ_2 then it is also CR wrt. their conjunction $\varphi_1 \wedge \varphi_2$. Thus, it is sufficient to show that *Maj* is CR wrt. all clauses D_j to conclude that *Maj* is CR wrt. their conjunction and hence with IC. Recall that all clauses D_j have maximal size 2. The case of a clause of size 1 is easily solved. Suppose $D_k = p_{j_k}$ or $D_k = \neg p_{j_k}$. Since all individuals must be rational the profile will be unanimous on issue j_k, and thus the majority will behave accordingly on issue j_k, in accordance with the constraint D_k. Let us then focus on a clause IC $= \ell_j \vee \ell_k$, where ℓ_j and ℓ_k are two distinct literals, i.e., atoms or negated atoms. A paradoxical profile for the majority rule with respect to this integrity constraint features a first majority of individuals not satisfying literal ℓ_j, and a second majority of individuals not satisfying literal ℓ_k. By the pigeonhole principle these two majorities must have a non-empty intersection, i.e., there exists one individual that does not satisfy both literals ℓ_j and ℓ_k, but this is incompatible with the requirement that all individual ballots satisfy IC.

($\Rightarrow$) Let us now assume for the sake of contradiction that IC is not equivalent to a conjunction of clauses of maximal size 2. We will now build a paradoxical situation for the majority rule with respect to IC.

We need the following crucial definition: Call *minimally falsifying partial assignment* (mifap-assignment) for an integrity constraint IC an assignment to some of the propositional variables that cannot be extended to a satisfying assignment, although each of its proper subsets can. We now associate with each mifap-assignment ρ for IC 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 IC. The negation of C_ρ is hence a disjunction, with the property of being a minimal clause implied by IC. Such formulas are known in the literature on knowledge representation as the *prime implicates* of IC, and it is a known result that every propositional formula is equivalent to the conjunction of its prime implicates (see, e.g., Marquis, 2000). Thus, we can represent IC with the equivalent formula $\bigwedge_\rho \neg C_\rho$ of all mifap-assignments ρ for IC. From our initial assumption we can infer that at least one mifap-assignment ρ^* has size > 2, for otherwise IC would be equivalent to a conjunction of 2-clauses.

We are now ready to show a paradoxical situation for the majority rule with respect to IC. Consider the following profile. Let y_1, y_2, y_3 be three propositional variables that are fixed by ρ^*. 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}. Recall that there are at least 3 individuals in $\mathcal{N}$. If there are other individuals, 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 2.5. In this profile, which can easily be 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. □

	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 2.8: A general paradox for the majority rule wrt. a clause of size 3.

Exercise 2.4. *What happens when the number of individuals is even? The majority rule does not have a unique definition: the strict majority accept an issue if $|N_j^{\boldsymbol{B}}| > \frac{n}{2}$ and the weak majority instead if $|N_j^{\boldsymbol{B}}| \geqslant \frac{n}{2}$. Does Theorem 2.1 still hold with these definitions?*

2.6 More Paradoxes! (Extra material)

In this section we describe two further paradoxes that can be analysed using the 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.

2.6.1 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).

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 2.9.

	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 2.9: The paradox of divided government.

As shown in Table 2.9, 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 the Exercise 2.1. 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.

2.6.2 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 2.10.

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 2.10.

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

Table 2.10: The multiple election paradox (MEP).

The outcome of the majority rule in Table 2.10 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 2.10 as a situation in which the three individual ballots are bound for instance by a budget constraint $\neg(p_A \wedge p_B \wedge p_C)$. 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 2.3. 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.[4] On the other hand, every instance of Definition 2.3 in binary aggregation represents an instance of the MEP, as the irrational outcome cannot have been voted for by any of the individuals.

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

[4] 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.

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).

Chapter 3

Collective Rationality and Languages for Integrity Constraints

Individual agents may be considered rational in many different ways, and for most cases it is possible to devise paradoxical situations leading to an irrational collective outcome (Lecture 2) The purpose of this lecture 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.

3.1 Two Definitions for Classes of Aggregation Procedures

3.1.1 Axiomatic definitions

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). We now present a list of 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, 2010), adapted to the framework of binary aggregation with integrity constraints.

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.

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.[1] 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) (recall from Lecture 1). 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.

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

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

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.

There are also other axioms, but for the moment we have enough definitions. 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 can finally define a notation to identify procedures that satisfy an axiom on the subdomain $\mathrm{Mod}(\mathrm{IC})^{\mathcal{N}}$ induced by a given integrity constraint IC.

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

Definition 3.1. *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* $\mathrm{IC} \in \mathcal{L}$ *the restriction $F_{\restriction \mathrm{Mod}(\mathrm{IC})^{\mathcal{N}}}$ satisfies the axioms in* AX*. This defines the following class:*

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

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.

We can show some properties of axiomatic classes of procedures. We write $\mathcal{F}[\mathrm{AX}]$ as a shorthand for $\mathcal{F}_{\{\top\}}[\mathrm{AX}]$, the class of procedures that satisfy the axioms in AX over the full domain $\mathcal{D}$.

Lemma 3.1. *The following facts hold:*
(*i*) *if $\mathcal{L}_1 \subseteq \mathcal{L}_2$ then $\mathcal{F}_{\mathcal{L}_1}[\mathrm{AX}] \supseteq \mathcal{F}_{\mathcal{L}_2}[\mathrm{AX}]$;*
(*ii*) *in particular, if $\top \in \mathcal{L}$, then $\mathcal{F}[\mathrm{AX}] \supseteq \mathcal{F}_{\mathcal{L}}[\mathrm{AX}]$;*
(*iii*) $\mathcal{F}_{\mathcal{L}}[\mathrm{AX}_1, \mathrm{AX}_2] = \mathcal{F}_{\mathcal{L}}[\mathrm{AX}_1] \cap \mathcal{F}_{\mathcal{L}}[\mathrm{AX}_2]$.

Exercise 3.1. *Prove Lemma 3.1.*

3.1.2 Definition in terms of collective rationality

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 $\mathcal{I}$.

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 the previous lectures we called an aggregation procedure collectively rational with respect to a formula $\mathrm{IC} \in \mathcal{L}_{PS}$ 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 3.2. *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 $\mathrm{IC} \in \mathcal{L}$:*

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

3.1.3 Collective Rationality and Languages

We now 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 3.2. *$\mathcal{CR}[\mathcal{L}^{\wedge}] = \mathcal{CR}[\mathcal{L}]$ 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., $\mathcal{CR}[\textit{cubes}] = \mathcal{CR}[\textit{literals}]$, 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., $\mathcal{CR}[clauses] = \mathcal{CR}[\mathcal{L}_{PS}]$. 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.

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.[2] Recall that two formulas are logically equivalent when they share 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 3.3. $\mathcal{CR}[\mathcal{L}^{\equiv}] = \mathcal{CR}[\mathcal{L}]$ *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. Bringing together the results of Lemma 3.2 and of Lemma 3.3, we can now give the following definition:

Definition 3.3. *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.

[2] 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.

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

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.

Lemma 3.4. $\mathcal{CR}[\mathcal{L} \cup \{\top\}] = \mathcal{CR}[\mathcal{L} \cup \{\bot\}] = \mathcal{CR}[\mathcal{L}]$ *for all* $\mathcal{L} \subseteq \mathcal{L}_{PS}$*.*

Exercise 3.2. *(For the more logic-oriented readers, arguably not very interesting) Does Definition 3.3 contain all the operations that we can perform on* $\mathcal{L}$ *that leave the set* $\mathcal{CR}[\mathcal{L}]$ *invariant? Yes. Show that given two languages for integrity constraints* $\mathcal{L}_1$ *and* $\mathcal{L}_2$*, if it is the case that* $\mathcal{L}_1 \neq \mathcal{L}_2$*, then* $\mathcal{CR}[\mathcal{L}_1] \neq \mathcal{CR}[\mathcal{L}_2]$ *(there is a detail missing in this statement).*

We conclude this section by establishing some easy properties of $\mathcal{CR}[\mathcal{L}]$:

Lemma 3.5. *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* $\mathcal{L}_1, \mathcal{L}_2 \subseteq \mathcal{L}_{PS}$*.*

Exercise 3.3. *Prove Lemma 3.5.*

3.1.4 From collective rationality to integrity constraints and back (extra material)

In the first part of the lecture 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 $\mathcal{I}$. We now introduce the inverse operation:

Definition 3.4. *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* $F \in \mathcal{G}$*:*

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

$\mathcal{LF}[\mathcal{G}]$ is the intersection of all $\mathcal{LF}[\{F\}]$ for $F \in \mathcal{G}$. The characterisation of the paradoxical constraints for the majority rule shown in the previous section can be rephrased as $\mathcal{LF}[Maj] = \mathit{2clauses}$.

Exercise 3.4. *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* $\mathcal{I}$*. Show that the following holds:*
(*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.*

3.2 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 3.1: 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.

3.2.1 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 3.1. $\mathcal{CR}[\mathit{literals}] = \mathcal{F}_{\mathit{literals}}[\text{U}]$.

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 3.2 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. $\square$

As remarked in Section 3.1.3, 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 3.1. $\mathcal{CR}[\mathit{cubes}] = \mathcal{F}_{\mathit{cubes}}[\text{U}]$.

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

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 3.2 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. $\square$

Exercise 3.5. *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 $\{p_j \leftrightarrow \neg p_k \mid p_j, p_k \in PS\}$. Show that $\mathcal{CR}[\mathcal{L}_{XOR}] = \mathcal{F}_{\mathcal{L}_{XOR}}[\mathrm{N}^{\mathcal{D}}]$.*

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 $\mathcal{F} = \{F : \mathcal{D}^{\mathcal{N}} \to \mathcal{D}\}$ is the class of all aggregation procedures (for fixed $\mathcal{I}$). We have already stated in Lemma 3.4 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 3.1. $\mathcal{CR}[\{\top\}] = \mathcal{CR}[\{\bot\}] = \mathcal{F}$.

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.[4]

Definition 3.5. *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 3.3. $\mathcal{CR}[\mathcal{L}_{PS}] = \mathrm{GDIC}$.

[4]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 (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).

Proof. Clearly, every generalised dictatorship lifts any arbitrary integrity constraint IC $\in \mathcal{L}_{PS}$. To prove the other direction, suppose that $F \not\in$ 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, $B_i \models$ IC for every $i \in \mathcal{N}$, i.e., $\boldsymbol{B}$ is a rational profile for the integrity constraint IC. But by construction, $F(\boldsymbol{B}) \not\models$ 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 $\mathcal{CR}[\mathcal{L}_{PS}]$. $\square$

Generalised dictatorships do not only include "bad" aggregation procedures. In Lecture 5 we will see the definition of some very interesting rules based on the selection of the most representative voters.

3.3 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.

3.3.1 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 3.1.3, for all the axiomatic properties considered we know that $\mathcal{F}[\text{AX}] \subseteq \mathcal{F}_{\mathcal{L}}[\text{AX}]$ 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 3.2. *The following equivalences hold:*
(*i*) $\mathcal{F}[\text{U}] = \mathcal{CR}[\textit{literals}]$.
(*ii*) $\mathcal{F}[\text{N}^{\mathcal{I}}] = \mathcal{CR}[\mathcal{L}_{\leftrightarrow}]$.

Proof. Refer to Theorems 3.1, 3.2. For all three classes we prove that $\mathcal{F}_{\mathcal{L}}[\text{AX}] = \mathcal{F}[\text{AX}]$, 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 $\boldsymbol{B} \in \text{Mod}(p_j)$ ($\boldsymbol{B} \in \text{Mod}(\neg p_j)$, respectively). Thus, all profiles in the scope of the axiom of unanimity are models of a literal. The other case is similar: 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. Therefore, for a procedure to satisfy unanimity or issue-neutrality on domains defined by $\mathcal{L}$ is equivalent to satisfying the same axiom on the full domain. □

Exercise 3.6. *Following up from Exercise 3.5, show that* $\mathcal{F}[\text{N}^{\mathcal{D}}] = \mathcal{CR}[\mathcal{L}_{XOR}]$.

3.3.2 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.

Proposition 3.2. *There is no language* $\mathcal{L} \subseteq \mathcal{L}_{PS}$ *such that* $\mathcal{CR}[\mathcal{L}] = \mathcal{F}_{\mathcal{L}}[\text{I}]$.

Proof. We prove this proposition by constructing, for any choice of a language $\mathcal{L}$, an independent function that is not collectively rational for a certain IC $\in \mathcal{L}$. Fix a language $\mathcal{L}$. This language will contain a falsifiable formula φ (otherwise $\mathcal{CR}[\mathcal{L}]$=$\mathcal{F}$ by Proposition 3.1 and we are done, as $\mathcal{F} \neq \mathcal{F}_{\mathcal{L}}[\text{I}]$). Choose a ballot/model $B^\star \in \mathcal{D}$ such that $B^\star \not\models \varphi$. Then the constant function $F \equiv B^\star$ is an independent function (on the full domain) that is not collectively rational. □

Proposition 3.3. *There is no language* $\mathcal{L} \subseteq \mathcal{L}_{PS}$ *such that* $\mathcal{CR}[\mathcal{L}] = \mathcal{F}_{\mathcal{L}}[\text{A}]$.

Proof. Same technique here. Fix a language $\mathcal{L}$. This language will contain a falsifiable formula φ (otherwise $\mathcal{CR}[\mathcal{L}]$=$\mathcal{F}$ by Proposition 3.1 and we are done, as $\mathcal{F} \neq \mathcal{F}_{\mathcal{L}}[\text{A}]$). Choose a ballot/model $B^\star \in \mathcal{D}$ such that $B^\star \not\models \varphi$. Then the constant function $F \equiv B^\star$ is an anonymous function (on the full domain) that is not collectively rational. □

Our interest toward these axioms does not cease here. On the contrary, the previous two propositions showed that the classes of independent and anonymous procedures behave in the same way as the full class $\mathcal{F}$ for what concerns collective rationality: the set of integrity constraints that are lifted by any aggregator in these classes is the trivial language $\{\top, \bot\}$. This suggests that interesting characterisations can be studied *inside* those classes, replacing the set $\mathcal{F}$ of all procedures with, e.g., the class $\mathcal{F}[\text{I}]$. This is an approach that has been studied widely, and we are going to pursue this approach in the fourth lecture.

3.3.3 Towards a general characterisation of axioms (extra material)

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,[5] but we may put forward the following informal definition.

Definition 3.6. *(*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 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 and independence 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

[5]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, 2013b).

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 3.3 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.

Exercise 3.7. *Give a formal definition of intra-profile and inter-profile axioms for aggregation procedures. Show formally that if an axiom can be characterised in terms of collective rationality then it is an intra-profile axiom. Show that inter-profile axioms canon be characterised in terms of collective rationality. Provide a characterisation of the set of intra-profile axioms that can be characterised in terms of collective rationality. (open problems, **prize**: an ice-cream for the first statement, large beer for each subsequent statement.)*

3.4 References

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). 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 (see Lecture 4).

Dokow and Holzman (2009, 2010) 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.

A similar approach has been taken by Nehring and Puppe (2010), who focused on the study of monotonic and independent procedures. The characterisation of paradoxical results for the majority rule presented in Lecture 2 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.

Characterisation results in line with those seen in this chapter relates the class of *quota rules*, i.e., aggregators that accept an issue if the ratio of acceptance exceed a given quota, and *languages of clauses*, i.e., disjunctions of limited size Grandi and Endriss (2013a). Many of these results are analogous to those proven by Dietrich and List (2007a) in the framework of (formula-based) judgment aggregation. From a computational perspective, however, the use of integrity constraints to model rationality assumptions, rather than referring to the consistency of judgment sets, leads to problems that are substantially easier to compute (see Lecture 1).

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.

Chapter 4

Ultrafilter methods

The aim of this lecture is to learn a widely used technique used to characterise classes of aggregation procedures. This technique is used when the axiom of independence is assumed, and focuses on the structure of coalitions of individuals and their power on the collective decision. We will first learn the technique in preference aggregation, presenting a proof of Arrow's Theorem and some related results, and then view another application of the same technique in (formula-based) judgment aggregation.

4.1 Arrow's Theorem and Winning Coalitions

The first use of the structure of winning coalitions to prove, in this case, a possibility result is due to Fishburn (1970). He showed that Arrow's Theorem does not generalise to the case of an infinite amount of individuals, by constructing an independent, unanimous and non-dictatorial procedure making use of an *ultrafilter* over individuals. His crucial observation was that the axiom of independence allows a rule to be defined in terms of *winning coalitions*, and hence the power of a set-theoretic formulation can be exploited. His work was later generalized to full characterisation theorems by Kirman and Sondermann (1972), and now most proofs in aggregation theory use similar techniques.

In this section we will prove Arrow's Theorem and related results using the so-called ultrafilter technique, showing its flexibility in proving (im)possibility theorems.

4.1.1 Reminder on Preference Aggregation

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. 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* $\boldsymbol{P} = (P_1, \dots, P_n)$. 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 $\boldsymbol{R} = (R_1, \dots, R_n)$ 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 4.1. *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* $\boldsymbol{P} = (P_1, \dots, P_n) \in \mathcal{L}(\mathcal{X})^{\mathcal{N}}$ 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.

4.1.2 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. 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 $aP_ib \Leftrightarrow aP'_ib$ 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 $a\, w(\boldsymbol{P})\, b \Leftrightarrow c\, w(\boldsymbol{P})\, d$.

4.1.3 Let's work in binary aggregation, it's simpler

Recall the embedding of PA into BA with IC (Lecture 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 enables 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 4.1.2 their equivalent formulation for binary aggregation.

For instance, the axiom of independence of irrelevant alternatives (IIA) is usually called simply independence (I), where $\mathcal{X} \subseteq \mathcal{D}^{\mathcal{N}}$ is a subset of the set of profiles:

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 $F(\boldsymbol{B})_j = F(\boldsymbol{B}')_j$.

The Pareto condition corresponds to a weak version of the axiom of unanimity (U), restricted to the case of individuals agreeing on the acceptance of an issue:

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 $F(\boldsymbol{B})_j = x$.

For the case of weak orders the resulting axiom is even weaker (but sufficient to obtain deep impossibilities, see Section 4.1.5).

Exercise 4.1. *State formally the binary aggregation versions of all axioms presented in Section 4.1.2. Show that the binary aggregation version of weak pareto is weaker than the axiom of unanimity for binary aggregation.*

Exercise 4.2. *The axiom of neutrality translates to the following statement:*

Issue-Neutrality *($N^{\mathcal{I}}$): 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 $F(\boldsymbol{B})_j = F(\boldsymbol{B})_{j'}$.*

Show formally that the two axioms correspond. Check with the relevant literature whether there are other possible formulations of neutrality, both in preference and in binary aggregation, and discuss the relations between the different formulations.

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 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 and A 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 $\mathrm{IC}_{<}$.

4.1.4 Winning coalitions and ultrafilters

First, some further notation (recall that we switched to binary aggregation): $N_j^{\boldsymbol{B}}$ denote with $N_j^{\boldsymbol{B}} = \{i \in \mathcal{N} \mid b_{i,j} = 1\}$ the set of individuals accepting issue j in profile $\boldsymbol{B}$. Let us consider aggregation procedures that satisfy the axiom of independence, proving a representation result in terms of *winning coalitions*:[1]

Proposition 4.1. *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. $\square$

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

Corollary 4.1. *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 $N_j^{\boldsymbol{B}} \in \mathcal{W}$.*

As is well understood in SCT, impossibility theorems in preference aggregation heavily feed on the notion of independence.

We now give the following definition, that of an *ultrafilter* (see, e.g., Davey and Priestley, 2002b):

Definition 4.2. *An **ultrafilter** $\mathcal{W}$ on a set $\mathcal{N}$ is a collection of subsets of $\mathcal{N}$ satisfying the following three conditions:*
(i) $\emptyset \notin \mathcal{W}$
(ii) $C_1, C_2 \in \mathcal{W}$ *implies* $C_1 \cap C_2 \in \mathcal{W}$ *(closure under intersection)*
(iii) C *or* $\mathcal{N} \setminus C$ *is in* $\mathcal{W}$ *for any* $C \subseteq \mathcal{N}$ *(maximality)*

The ultrafilter technique in a nutshell: start from a class of independent aggregation procedures, show that the properties defining this class force the set of winning coalitions to be an ultrafilter, conclude by observing that an ultrafilter over a finite set is principal and hence that there is an individual, the dictator, that is included in every winning coalition.

[1]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).

4.1.5 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 provide an alternative proof of this theorem using the ultrafilter techniques and the binary aggregation translation of preferences. 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 (binary aggregation), 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 4.1. *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 4.1, 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 4.1.

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 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. $\square$

[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.

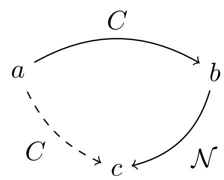

Figure 4.1: Collective transitivity entails issue-neutrality.

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 4.1 can be extended to these more restrictive constraints, obtaining a proof of what is known in PA as the *"contagion lemma"*:

Lemma 4.2. *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 4.1 (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 4.1.3, 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 4.2 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 4.1, we can characterise F_w in terms of the set of winning coalitions $\mathcal{W}$. We now prove that $\mathcal{W}$ is an *ultrafilter* (Definition 4.2). 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 4.2. 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 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 $\mathcal{W}$.

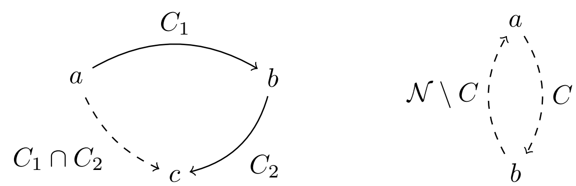

Figure 4.2: The set of winning coalitions is an ultrafilter.

(*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 4.2. 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. $\square$

Notice that in the proof of Theorem 4.1 we have not used the assumption of reflexivity of weak orders, and that Lemma 4.2 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 4.1 in place of Lemma 4.2, the same proof shows the following result:

Theorem 4.2. *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 4.1. 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 4.3. *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 4.1 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, 2002a). 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 4.3. 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 $D \in \mathcal{W}$.

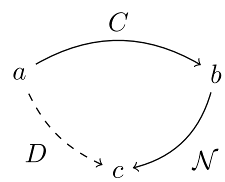

Figure 4.3: The set of winning coalitions is a filter.

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. $\square$

If you are wondering what Kirman and Sondermann (1972) said, here is a statement of their theorem:

Theorem 4.4. *Given a set of individuals $\mathcal{N}$ and a set of alternatives $\mathcal{X}$ such that $|\mathcal{X}| \geqslant 3$, the winning coalitions $\mathcal{W}$ associated to an independent and unanimous SWF for $\mathcal{N}$ and $\mathcal{X}$ form an ultrafilter.*

Arrow's Theorem is the corollary of Theorem 4.4 in case both sets of individuals and alternatives are finite.

Exercise 4.3. *Prove Theorem 4.4 (very easy!). Exhibit an independent and unanimous SWF that is non-dictatorial.*

4.2 Ultrafilter techniques in judgment aggregation

Proofs in the setting of judgment aggregation tend to make a heavy use of winning coalitions, although not always the ultrafilter technique explained in the previous section. In this section we prove one of the central theorems of judgment aggregation, due to Nehring and Puppe (2007), by making use of the ultrafilter technique.

Recall the framework of judgment aggregation. We have formulas in a propositional logical language, we have an agenda Φ which is closed under complementation and does not contain double negation, and we are interested in *consistent* aggregation, i.e., starting from satisfiable individual subsets of the agenda we want to obtain a satisfiable set of formulas in the output. As for axioms, it is sufficient to substitute $\varphi \in J_i$ for $b_{i,j} = 1$ in binary aggregation or aP_ib in preference aggregation to obtain their analogous in judgment aggregation.

We need the following definition, that formalizes a limit to the complexity of the formulas involved in the aggregation:

Definition 4.3. *An agenda Φ satisfies the median property iff all inconsistent subsets of Φ contain an inconsistent subset of size 2.*

An agenda that satisfies the median property has inconsistencies which are very simple, since they can only involve two formulas. Longer chain of inconsistent formulas that are consistent if taken in smaller subsets are not allowed. For instance, the agenda $\{p, q, p \wedge q\}$ closed under negation does not satisfy the median property: the inconsistent subset of size three $\{p, q, \neg(p \wedge q)\}$ is minimal, i.e., it does not contain any inconsistent subset. The agenda $\{p, r, p \wedge q\}$ closed under negation, on the other hand, satisfies the median property.

We need one last axioms, which we neglected throughout the course, that is a generalization of the positive responsiveness property by May (Lecture 1) and formalizes the intuition that if an accepted formula gets additional support it should still be accepted:

- **I-Monotonicity** (M^{I}): For any formula $\varphi \in \Phi$ and any two profiles $\boldsymbol{J}, \boldsymbol{J}' \in \mathcal{J}(\Phi)^{\mathcal{N}}$, 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 $\varphi \in F(\boldsymbol{J}')$

Let us finally state and prove the following result, due to Nehring and Puppe (2007):

Theorem 4.5. *Let $\mathcal{N}$ be of odd cardinality. An agenda Φ satisfies the median property if and only if there exists unanimous, neutral, independent, monotonic and non-dictatorial aggregation procedures for Φ that are complete and consistent.*

Proof. For the first direction we need to show that if the agenda Φ satisfies the median property then we can construct an aggregator that satisfies all desirable conditions. Well, just use the majority rule! (Exercise 4.4).

Let us now show by contradiction that if the agenda does not satisfy the median rule then every neutral, independent and monotonic aggregation rule is a dictatorship. Any idea of how to show this? Using the ultrafilter method!

First, we need to show that any independent and neutral aggregation procedure can be defined in terms of winning coalitions (by independence), and that this set is equal for each formula (by neutrality). No surprise here, proofs work as for the case of binary aggregation. Now we need to show that the set $\mathcal{W}$ of winning coalitions is an ultrafilter:

(i) $\emptyset \notin \mathcal{W}$: if $\emptyset \in \mathcal{W}$ then by monotonicity also $\mathcal{N} \in \mathcal{W}$ against consistency (φ and $\neg\varphi$ would always be accepted).

(ii) $\mathcal{W}$ is closed under intersection: since we assumed that Φ does not satisfy the median property, we know that there must exists an inconsistent subset $X \subseteq \Phi$ of size at least three, and that does not contain any inconsistent subset of size 2. Let φ_1, φ_2 and φ_3 be three distinct formulas in X. Let now C_1 and C_2 be winning coalitions, and consider the following profile:

$C_1 \setminus (C_1 \cap C_2)$	φ_1	$\neg\varphi_2$	φ_3	$X \setminus \varphi_1, \varphi_2, \varphi_3$
$C_2 \setminus (C_1 \cap C_2)$	$\neg\varphi_1$	φ_2	φ_3	$X \setminus \varphi_1, \varphi_2, \varphi_3$
$C_1 \cap C_2$	φ_1	φ_2	$\neg\varphi_3$	$X \setminus \varphi_1, \varphi_2, \varphi_3$
$\mathcal{N} \setminus (C_1 \cup C_2)$	$\neg\varphi_1$	φ_2	φ_3	$X \setminus \varphi_1, \varphi_2, \varphi_3$

Now this means that:

- $\varphi_1 \in F(\boldsymbol{J})$, since $N^{\boldsymbol{J}}_{\varphi_1} = C_1 \in \mathcal{W}$
- $\varphi_2 \in F(\boldsymbol{J})$ since $N^{\boldsymbol{J}}_{\varphi_2} \supseteq C_2$ and $C_2 \in \mathcal{W}$ hence by monotonicity also $N^{\boldsymbol{J}}_{\varphi_2} \in \mathcal{W}$.
- all formulas in $X \setminus \{\varphi_1, \varphi_2, \varphi_3\}$ are accepted since $\mathcal{N} \in \mathcal{W}$ (by completeness, since $\emptyset \notin \mathcal{W}$ then $\mathcal{N} \in \mathcal{W}$).
- this implies that $\neg\varphi_3 \in F(\boldsymbol{J})$.

To conclude it is sufficient to observe that $N^{\boldsymbol{J}}_{\neg\varphi_3} = C_1 \cap C_3$, hence $C_1 \cap C_3 \in \mathcal{W}$.

(iii) $\mathcal{W}$ is maximal. Let $W \subset \mathcal{N}$, and construct a profile in which only those individuals in W accept a formula φ. Observe that by individual rationality exactly those individuals in $\mathcal{N} \setminus W$ accept $\neg\varphi$. Now by completeness either φ is accepted by F, hence W is a winning coalitions, or $\neg\varphi$ is accepted, hence $\mathcal{N} \setminus W$ is a winning coalition.

It is now sufficient, as usual, to observe that an ultrafilter over a finite set is principal, to conclude that the rule must be dictatorial. □

The previous proof was obtained by means of axiomatic properties mostly, and those conditions can be relaxed obtaining different conditions on the agenda (the following section list some references where these results can be found).

Exercise 4.4. *Show that the majority rule with an odd number of individuals and an agenda that satisfies the median property is consistent, complete, anonymous, unanimous, independent and monotonic.*

Exercise 4.5. *Theorem 4.5 holds also for an even number of individuals, with the same statement. Can you prove it in that case? What is the most difficult direction?*

Exercise 4.6. *Exploiting the translation of judgment aggregation in binary aggregation that was sketched in Lecture 2, state the equivalent version of Theorem 4.5 in binary aggregation with integrity constraints. Does this statement remind you of something? What is the equivalent of the median property for integrity constraints?*

4.2.1 References

For a more detailed study of the use of ultrafilters in preference aggregation we refer to Daniëls and Pacuit (2009) and Herzberg and Eckert (2012). Judgement aggregation proofs are most often stated in terms of winning coalitions, for an explicit use of the ultrafilter techniques we refer to Grossi and Pigozzi (2014).

A very general formulation of Arrow's Theorem in the slightly more general setting of graph aggregation can be found in (Endriss and Grandi, 2014b). This is a very useful albeit technical reading to understand ultrafilter techniques in a very abstract setting.

Chapter 5

Aggregation procedures, computational complexity, and applications

We have seen many theorems in the last four lectures, but only one aggregation procedure! This procedure was the notorious majority rule, which in many cases did not even satisfy basic criteria like collective rationality.

In this lecture we present a long list of aggregation procedures in the setting of formula-based judgment aggregation, with the main aim to enforce collective rationality. We will briefly talk about what axioms are satisfied, but mostly we will assess aggregation procedures by studying the computational complexity of winner determination, i.e., how hard is it to compute the result of a given aggregator. We conclude by listing some applications of (mostly judgment and binary) aggregation to a number of diverse settings: crowdsourcing (collective annotations and sentiment analysis), strategic aspects of aggregation, multi-agent argumentation, modal logic and aggregation.

5.1 Winner determination and computational complexity

Recall briefly the setting of *formula-based judgment aggregation.* We have an agenda Φ, we collect individual judgment sets $J_i \subseteq \Phi$ which are complete and consistent, and we want to obtain a complete and consistent collective judgment set.

The *majority rule*, the only rule we have seen so far except for dictatorships and constant functions, accepts a given formula φ if and only if the set of individuals accepting φ in profile $\boldsymbol{J}$, i.e., $N_\varphi^{\boldsymbol{J}}$, has cardinality larger or equal than $\frac{n+1}{2}$, where n is the number (odd) of individuals.

Aside from axiomatic properties, which are a very good mean to assess the qualities of a given aggregator, our computationally oriented mind brings us to study how hard is it to use or implement a given rule. *Computational complexity*, a well-established topic in computer science, provides all the tools that we need for this task.[1]

But what does it mean for a rule to be complex to use? What is the exact algorithmic problem we need to implement when we use an aggregation procedure? We now give a precise formulation of this problem.

The problem of *winner determination* in voting theory is that of computing the election winner given a profile of preferences supplied by the voters. The corresponding decision problem asks, given a preference profile and a candidate, whether the given candidate is the winner of the election. In JA, we want to compute $F(\boldsymbol{J})$ for a given profile $\boldsymbol{J}$. For a resolute aggregation procedure F, we can formulate a corresponding decision problem by asking, for a given formula, whether it belongs to $F(\boldsymbol{J})$:

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

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

For the case of irresolute JA procedures F, i.e., procedures that output a *set* of collective judgment sets rather than a single one, we can adapt the winner determination problem in the following way:

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

To see that this is an appropriate formulation of 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 including an arbitrarily chosen first formula of the agenda φ_1, i.e., $L = \{\varphi_1\}$. In case the answer is positive, consider a second formula φ_2 and query WinDet* with $L = \{\varphi_1, \varphi_2\}$. Use subset $L = \{\sim\varphi_1, \varphi_2\}$ in case of a negative answer. Continue this process until all formulas in the agenda have been covered.

[1] A useful brief introduction to computational complexity can be found here: `http://staff.science.uva.nl/~ulle/teaching/comsoc/2009/comsoc-complexity.pdf`

5.2 Quota rules

An aggregation procedure F for $n = |\mathcal{N}|$ individuals is a *quota rule* if for every formula φ there exists a quota $q_\varphi \in \{0, \dots, n{+}1\}$ such that $\varphi \in F(\boldsymbol{J})$ if and only if $|N^{\boldsymbol{J}}_\varphi| \geqslant q_\varphi$. The class of quota rules has been studied in depth by Dietrich and List (2007a). Let us focus on a particular class of quota rules:

Definition 5.1. *Given some $m \in \{0, \dots, n{+}1\}$ and an agenda Φ, the* ***uniform quota rule*** *with quota m is the aggregation procedure F_m with $\varphi \in F_m(\boldsymbol{J}) \Leftrightarrow |N^{\boldsymbol{J}}_\varphi| \geqslant m$.*

An aggregation procedure satisfies A, I, $\mathrm{M^I}$, and N if and only if it is a uniform quota rule; this fact follows immediately from a result by Dietrich and List (2007a), who use a slightly more narrow definition of quota rule. Provided $m \neq n + 1$, the uniform quota rule F_m also satisfies U.

A quota rule of special interest is the *majority rule*. The majority rule is the uniform quota rule with $m = \frac{n+1}{2}$; it accepts a formula whenever there are more individuals accepting it than there are rejecting it (recall that we did assume n to be odd).

Exercise 5.1. *Show that the majority rule is the only uniform quota rule that always outputs a complete and complement-free judgment set (i.e., it does never accept a formula together with its negation).*

Depending on the particular thresholds defining a quota rule, there are specific requirements on the agenda to guarantee collective rationality. For the details see the work of Dietrich and List (2007a).

5.2.1 Winner determination for quota rules

It is immediately clear that winner determination is a polynomial problem for any quota rule, including the majority rule.

Lemma 5.1. $\textsc{WinDet}(F_m)$ *is in* P *for any uniform quota rule F_m.*

Exercise 5.2. *Show an algorithm that solves* $\textsc{WinDet}(F_m)$ *in polynomial time.*

5.3 The premise-based and conclusion-based procedures

Two basic aggregation procedures that can be set up in a way so as to avoid the problem of collective rationality have been discussed in the JA literature from the very beginning, 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. Under 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; 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 (so strictly speaking it does not conform to our definition of an aggregation procedure), and we shall not consider it 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.

For many JA problems, it may be natural to divide the agenda into premises and conclusions. Let $\Phi = \Phi^p \uplus \Phi^c$ 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 5.2. *The* ***premise-based procedure*** PBP *for* Φ^p *and* Φ^c *is the function mapping each profile* $\boldsymbol{J} = (J_1, \ldots, J_n) \in \mathcal{J}(\Phi)^n$ *to the following judgment set:*

$$\begin{array}{rcl} \mathrm{PBP}(\boldsymbol{J}) & = & \Delta \cup \{\varphi \in \Phi^c \mid \Delta \models \varphi\}, \\ & & \textit{where } \Delta = \{\varphi \in \Phi^p \mid |N_\varphi^{\boldsymbol{J}}| \geqslant \frac{n+1}{2}\} \end{array}$$

That is, Δ is the set of premises accepted by a (strict) majority; and the PBP will return this set Δ together with those conclusions φ that logically follow from Δ ($\Delta \models \varphi$).

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*, then we have to impose restrictions on the premises. We have seen in the previous lecture that 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 $\leqslant 2$. 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*, then 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.

Here we make two additional assumptions. 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 under these assumptions.

So, to summarise, the instance of the PBP we shall work with 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 individuals 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 input profiles, and assuming that n is odd, this leads to a resolute JA procedure that is consistent and complete. On the downside, the PBP violates most of the standard axioms typically considered, such as N and I.

Exercise 5.3. *Show that the PBP violates U (unanimity).*

5.3.1 Winner determination for the premise-based procedure

Winner determination is also tractable for the premise-based procedure:

Proposition 5.1. 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. $\square$

5.4 The distance-based procedure

The basic idea of a distance-based approach to aggregation is to select an outcome that, in some sense, minimises the distance to the input profile. This idea has been used extensively in both preference aggregation (Kemeny, 1959) and belief merging (Konieczny and Pino Pérez, 2002). The first example of a JA procedure based on a notion of distance was introduced by Pigozzi (2006), albeit under the restrictive assumption that the agenda is closed under propositional variables and that each compound formula will either be unanimously accepted or unanimously rejected by all agents. Most importantly, in Pigozzi's approach 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 has later been defined by Miller and Osherson (2009), which these authors call the *Prototype-Hamming* rule. This is the distance-based procedure we shall define and analyse here. It is an *irresolute* procedure, returning a (nonempty) set of collective judgment sets.

First, let us define a notion of distance between judgment sets. We can interpret a judgment set J as a (characteristic) function $J : \Phi \to \{0, 1\}$ with $J(\varphi) = 1$ if $\varphi \in J$ and $J(\varphi) = 0$ if $\varphi \notin J$. The *Hamming distance* $H(J, J')$ between two (complete and complement-free) judgment sets J and J' is the number of positive formulas on which they differ:

$$H(J, J') \quad = \quad \sum_{\varphi \in \Phi^+} |J(\varphi) - J'(\varphi)|$$

Definition 5.3. *Given an agenda Φ, the* ***distance-based procedure*** DBP *is the function mapping each profile $\boldsymbol{J} = (J_1, \ldots, J_n) \in \mathcal{J}(\Phi)^n$ to the following set of judgment sets:*

$$\mathrm{DBP}(\boldsymbol{J}) \quad = \quad \operatorname*{argmin}_{J \in \mathcal{J}(\Phi)} \sum_{i \in \mathcal{N}} H(J, J_i)$$

A collective judgment set under the DBP minimises the amount of disagreement with the individual judgment sets (i.e., it minimises the sum of the Hamming distances with all individual judgment sets). Note that in cases where the majority rule leads to a consistent outcome, the outcome of the DBP coincides with that of the majority rule (making it a resolute procedure over these profiles). We can combine the DBP with a *tie-breaking rule* to obtain a resolute procedure.

The DBP is complete and consistent by design: only judgment sets in $\mathcal{J}(\Phi)$ are considered candidates when searching for a solution. However, it violates most of the standard axiomatic properties when those are adapted to the case of irresolute JA procedures (Lang et al., 2011). In particular, the DBP is not independent; indeed, it is based on the very idea that correlations between propositions should be exploited rather than neglected.

5.4.1 Winner determination of the distance-based procedure (extra material)

We now want to analyse the complexity of the winner determination problem for the distance-based procedure. As the DBP is irresolute, we study the decision problem WinDet*. As we shall see, WinDet*(DBP) is Θ_2^p-complete, thus showing that this rule is very hard to compute. The class Θ_2^p (also known as $\Delta_2^p(O(\log n))$, $\mathrm{P}^{\mathrm{NP}[\log]}$ 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 or, equivalently, that can be solved in polynomial time asking a polynomial number of such queries in parallel (Wagner, 1987; Hemachandra, 1989). 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:

WinDet$_K^\star$(DBP)
Instance: Agenda Φ, profile $\boldsymbol{J} \in \mathcal{J}(\Phi)^n$, subset $L \subseteq \Phi$, $K \in \mathbb{N}$.
Question: Is there a $J^\star \in \mathcal{J}(\Phi)$ with $L \subseteq J^\star$ such that $\sum_{i \in \mathcal{N}} H(J^\star, J_i) \leqslant K$?

That is, we ask whether there exists a judgment set $J^\star$ with a Hamming distance to the profile of at most K that accepts all the formulas in L. In other words, rather than aiming at computing a winning judgment set, this problem merely allows us to compute a judgment set of a certain minimal quality (where quality is measured in terms of the Hamming distance). It is easy to show that this problem lies in NP.

Lemma 5.2. WinDet$_K^\star$(DBP) *is in* NP.

Proof. NP is composed by problems that can be solved by first guessing certificate in an exponential space and then check in polynomial time whether the certificate solves positively the problem posed. We therefore devise a non-deterministic polynomial algorithm to solve the problem of $\text{WinDet}^{\star}_{K}$(DBP).

Recall that we are given an agenda Φ, a profile $\boldsymbol{J} \in \mathcal{J}(\Phi)^n$, a subset $L \subseteq \Phi$, and a number $K \in \mathbb{N}$. What we want to know is whether there exists a $J^\star \in \mathcal{J}(\Phi)$ with $L \subseteq J^\star$ that is at most K-distant to the profile.

To solve the problem it is sufficient to first guess a complete judgment set $J^\star$ such that $L \subseteq J^\star$ and, *at the same time*, an assignment ρ to all variables contained in Φ. The space of all such profiles and assignments is exponential in the size of the input. Now it is sufficient to check that: (i) $\rho \models J^\star$, hence $J^\star$ is consistent, and (ii) that $\sum_{i \in \mathcal{N}} H(J^\star, J_i) \leqslant K$, to obtain an answer to $\text{WinDet}^{\star}_{K}$(DBP). □

To obtain an upper bound for the winner determination problem for the DBP, we can now use a standard construction. This first involves identifying the "best" value for K, and then deciding $\text{WinDet}^{\star}_{K}$(DBP) for that value of K. The latter can be done with a logarithmic number of queries to the problem the complexity of which we have analysed in Lemma 5.2. Together, this yields the desired upper bound:

Lemma 5.3. $\text{WinDet}^{\star}$(DBP) *is in* Θ^p_2.

Proof. The problem $\text{WinDet}^{\star}$(DBP) asks whether there exists a *winning* judgment set that accepts all formulas in a given subset $L \subseteq \Phi$. Since the Hamming distance between a judgment set and a profile is bounded from above by a polynomial figure, we can solve this problem by performing a binary search over K using a logarithmic number of queries to $\text{WinDet}^{\star}_{K}$(DBP).

More precisely, since $\sum_{i \in \mathcal{N}} H(J^\star, J_i)) \leqslant K^\star = \frac{|\Phi|}{2} \cdot |\mathcal{N}|$, a figure that is polynomial in the size of the problem description, we can ask a first query to $\text{WinDet}^{\star}_{K}$(DBP) with $K = \frac{K^\star}{2}$ and an empty subset of designated formulas. 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 $K = \frac{3}{4} \cdot K^\star$. This process 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 subset L 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 judgment set with Hamming distance strictly less than K^w, one of the winning judgment sets contains all formulas in L. On the other hand, in case of a negative answer all judgment sets containing L have Hamming distance bigger than K^w, and thus cannot belong to the winning set. □

It is also possible to show that the upper bound established by Lemma 5.3 is tight. We refer to Endriss et al. (2012) for the details.

5.5 Rules based on minimization

There are many alternative rules that can be defined based on minimisation principles. The following examples are taken from the work of Lang et al. (2011), and are all *majority preserving* rules, i.e., they output the result of the majority rule in case it is consistent.

Exercise 5.4. *Show that the DBP is majority consistent.*

Let $Maj(\boldsymbol{J})$ be the result of the majority rule over a profile $\boldsymbol{J}$: let $\mathrm{Maj}_{\subseteq}(\boldsymbol{J})$ be the set of all maximal consistent subsets of $Maj(\boldsymbol{J})$ by inclusion, and $\mathrm{Maj}_{|.|}(\boldsymbol{J})$ be the set of all consistent subsets of $Maj(\boldsymbol{J})$ which have maximal cardinality.

Definition 5.4. *The* maximal sub-agenda rule *associates with every profile those judgment sets that extends a maximal consistent subset of the outcome of the majority rule:*

$$\mathrm{MSA}(\boldsymbol{J}) = \{J \in \mathcal{J}(\Phi) \mid J \supseteq M \ \textit{with}\ M \in Maj_{\subseteq}(\boldsymbol{J})\}\}$$

Definition 5.5. *The* maxcard sub-agenda rule *associates with every profile those judgment sets that extends a consistent subset of maximal cardinality of the outcome of the majority rule:*

$$\mathrm{MCSA}(\boldsymbol{J}) = \{J \in \mathcal{J}(\Phi) \mid J \supseteq M \ \textit{with}\ M \in Maj_{|.|}(\boldsymbol{J})\}\}$$

Instead of modifying directly the outcome of the majority rule in a minimal way so that it can be consistent, another approach is that of modifying the profile of individual judgment sets in a minimal way so that the outcome of the majority over this new modified profile is consistent. If $\boldsymbol{J}$ and $\boldsymbol{J}'$ are two profiles, let $\mathcal{H}(\boldsymbol{J}, \boldsymbol{J}') = \sum_{i \in \mathcal{N}} H(J_i, J'_i)$.

Definition 5.6. *The* minimal number of atomic changes *rule associates with every profile the closest consistent outcome of the majority rule:*

$$\begin{aligned}\mathrm{MNAC}(\boldsymbol{J}) = \{\mathrm{Maj}(\boldsymbol{J}') \mid \mathrm{Maj}(\boldsymbol{J}') \in \mathcal{J}(\Phi) \ \textit{and}\ \mathcal{H}(\boldsymbol{J}, \boldsymbol{J}') \leqslant \mathcal{H}(\boldsymbol{J}, \boldsymbol{J}'') \\ \textit{for all}\ \boldsymbol{J}'' \in \mathcal{J}(\Phi)^{\mathcal{N}}\}\end{aligned}$$

Another possibility is to start using the weights of acceptance/rejection of the formulas in the agenda. The DBP does something similar in minimizing the overall weights over all formulas. Inspired by the work of Tideman (1987) in voting theory we can give the following definition.

First, we need some notation. For a given profile $\boldsymbol{J}$, define the *majority strength* of a formula φ as $\mathrm{MS}^{\boldsymbol{J}}(\varphi) = \max\{|N^{\boldsymbol{J}}_{\varphi}|, |N^{\boldsymbol{J}}_{\neg\varphi}|\}$, inducing an ordering $\succ^{\boldsymbol{J}}_{\tau}$ on formulas, with ties broken using a permutation $\tau : \Phi^+ \to \Phi^+$, where Φ^+ is the set of positive formulas in the agenda. Now, for a given profile $\boldsymbol{J}$ and permutation $\tau : \Phi^+ \to \Phi^+$, we define the judgment set $L^{\boldsymbol{J}}_{\tau}$ via the following procedure:

for $\varphi \in \Phi$, following order $\succ_\tau^{\boldsymbol{J}}$ **do**
$L = L \cup \varphi$ if $L \cup \varphi$ is consistent
$L = L \cup \neg\varphi$ otherwise

That is, we go through the formulas in order of majority strength (with ties broken by τ) and take over the majority decision whenever this is consistent with the previous choices, otherwise we invert it.

Definition 5.7. *The* ***ranked-agenda rule*** *is the aggregation rule that selects the individual ballots returned by the above procedure for some tie-breaking rule τ:*

$$\mathrm{RA}(\boldsymbol{J}) \quad = \quad \{L_\tau^{\boldsymbol{J}} \mid \tau \textit{ is a permutation on } \Phi^+\}$$

The computational complexity of winner determination for all these rules has been studied by Lang and Slavkovik (2014), albeit for a different algorithmic problem than the one presented here. The results are discouraging: for all cases the problem sits either in the second layer of the polynomial hierarchy or it is in Θ_2^p, as for the DBP. Using rules based on minimization is hard!

5.6 Rules based on the most representative voter

Lemma 5.3 showed that computing the result of the distance-based procedure is very hard. But, on the other hand, the distance-based rule seems to find the ideal compromise among all judgment sets.[2] A classic solution to this problem is to resort to *polynomial approximations*, i.e., polynomial algorithms that compute a judgment set which is very close to the ideal result of the DBP.

This approach has been pursued extensively in preference aggregation, where the DBP is called the Kemeny rule (Dwork et al., 2001; Ailon et al., 2008; Kenyon-Mathieu and Schudy, 2007). However, similar algorithms are usually randomized and are mostly focused on an efficient computation and hence cannot be explained and used as practical aggregation rules themselves.

A different possibility is that of restricting the search space of a distance-based procedure to those judgment sets that were *received* from the individuals. Consider the following two definitions:

Definition 5.8. *The* ***average-voter rule*** *is the aggregation rule that selects those individual judgment sets that minimise the sum of the Hamming distances to the individual judgment sets:*

$$\mathrm{AVR}(\boldsymbol{J}) \quad = \quad \underset{J \in \textsc{Supp}(\boldsymbol{J})}{\operatorname{argmin}} \sum_{i \in \mathcal{N}} H(J, J_i)$$

[2]With some doubts on the fact that the Hamming distance treats all formulas in the same way, although this can be solved by considering particular distances created for judgment sets, such as the one developed by Duddy and Piggins (2012).

Definition 5.9. *The* ***majority-voter rule*** *is the aggregation rule that selects those individual judgment sets that minimise the Hamming distance to one of the majority outcomes:*

$$\mathrm{MVR}(\boldsymbol{J}) = \operatorname*{argmin}_{J \in \textsc{Supp}(\boldsymbol{J})} H(J, \mathrm{Maj}(\boldsymbol{J}))$$

Similar rules are clearly polynomial to compute: it is sufficient to read the input profile and compute the distances of each individual judgment sets to the others, or to the outcome of the majority rule. Moreover, as we showed in previous work Endriss and Grandi (2014a), they are very good approximations of the distance based procedure.[3]

5.6.1 References

The first paper to study the complexity of judgment aggregation procedures is the work of Endriss et al. (2012), followed up by Baumeister et al. (2011, 2013) for the study of strategic issues for the premise-based procedure. Many aggregation procedures based on minimization were introduced by Lang et al. (2011), and the complexity of winner determination for many such rules has been studied in Lang and Slavkovik (2014). Judgment aggregation procedures based on a notion of score (like the Borda rule in voting) have been defined by Dietrich (2014). The relations between judgment aggregation rules and well-studied voting procedures have been spelled out by Lang and Slavkovik (2013).

5.7 Applications

Here is a selection of current research topics where judgment/binary aggregation is being used, with references for further readings.

5.7.1 Multiagent argumentation

The problem of (abstract) argumentation is that of selecting a "winning" argument among a set of possibly conflicting arguments. More formally, we are given an *attack graph* among arguments, and we need to select a subset of arguments that are in some ways acceptable, called an *extension* or a *labeling*. There are many possible ways to do so, depending on the characteristics we want on the extension. For instance, we may require that no pair of arguments in the extension attack each other (conflict-free).

When multiple agents are involved, we may face two problems:

- we may observe different extensions from different agents on the same attack graph.
- we may observe different attack graphs.

[3] This is true in binary aggregation with integrity constraints, but it is very likely that the results extend to the formula-based framework for judgment aggregation used above.

In the first case we need to aggregate extensions over the same graph, making sure that we obtain an acceptable extension in the output (sounds like collective rationality...). Rahwan and Tohmé (2010), as well as Caminada and Pigozzi (2011), have worked on this problem using techniques from (formula-based) judgment aggregation.

In the second case we are simply aggregating graphs. Coste-Marquis et al. (2007) have framed on this problem in the setting of belief merging.[4]. Endriss and Grandi (2014b) have worked on the problem of graph aggregation from a more general perspective not focused on the problem of argumentation .

5.7.2 Crowdsourcing: collective annotation of corpora and sentiment analysis

In a series of papers, Endriss and Fernández (2013); Qing et al. (2014); Kruger et al. (2014) have used settings related to binary aggregation to study an important problem in computational linguistics, that of *collective annotations*. This is a crowdsourcing problem: a set of words or phrases that needs to be classified to create an annotated corpus is sent out to a number of people that provide their classification, and this information needs to be aggregated to create the collective annotation. This is used for instance in computational linguistics to create large corpora of annotated phrases. They were successful in showing that their aggregation-based methods outperform the standard ones, opening up a very fruitful area of application for aggregation methods.

Another problem that falls under the crowdsourcing label is that of *sentiment analysis*: assume that we want to know which is the most popular candidate in an election, or the most preferred products in a catalogue; after collecting a large number of textual expressions (tweets, blog posts, reviews) from internet we classify them by making use of automatic classifiers into positive and negative, and then we need to aggregate this information into a collective sentiment that allows us to detect what is the most preferred candidate/alternative. A position paper by Grandi et al. (2014b) presents some possible research direction for the use of aggregation methods in this area. It is mostly focused on the use of voting rules or preference aggregation procedures, but we can easily imagined to be able to extract more complex information from individual text and make the use of binary or judgment aggregation necessary.

5.7.3 Aggregation and modal logic

If we can aggregate graphs, then why not studying constraints in modal logic? If we can aggregate propositional formulas, why not aggregating modal formulas?

Formula-based judgment aggregation have been studied in general logics (i.e., any logic that satisfies some criteria including monotonicity, which does not rule out most modal logics) by Dietrich (2007). No specific study of modal logic has been done to date.

[4]The relations between belief merging and judgment/binary aggregation are subject of ongoing research

The aggregation of Kripke models using integrity constraints formulated in modal logic is an interesting research topic (on which I am currently working on with Ulle Endriss). A nice property of this setting is that collective rationality does not have a single definition. We can interpret the integrity constraint:

- on specific worlds in the models (truth at a given world)
- on all worlds in the models (global truth)
- on all models on all worlds (validity)

Observe that the last possible definition corresponds to collective rationality with respect to a given graph property, using the well-known results in correspondence theory in modal logic.

5.7.4 Strategic aspects

When a collectivity needs to take a decision there are personal interest involved: individuals can misrepresent their true opinion/preference/judgment to obtain a result they favor. This problem is called *strategic voting* or *manipulation*, and it is widely studied in preference aggregation and in voting theory.

Judgment aggregation does not provide us with enough information on the individual preferences to study the problem of manipulation directly. We simply receive a judgment set, and we have make some assumptions on which collective judgments will be preferred to others based on this information:

- we can assume that an individual wants to get her complete judgment set as the collective outcome. In this case manipulation is most often impossible.
- we can assume that individual wants to get one particular formula in the collective judgment set. This case is studied by Dietrich and List (2007c), where it is proven that manipulation is impossible iff the aggregation procedure is independent and monotonic.
- we can assume that an individual wants to get as many formulas as possible coherent with her judgment set. In this case we induce an individual preference over judgment sets using the Hamming distance. This case has also been studied by Dietrich and List (2007c), and its computational aspects by Endriss et al. (2012); Baumeister et al. (2011, 2013).

A novel research trend in strategic voting is the use of game-theoretic notion to characterize equilibria of manipulation. A first study in the setting of binary aggregation has been done by Grandi et al. (2014a).

Bibliography

T. Ågotnes, W. van der Hoek, and M. Wooldridge. On the logic of preference and judgment aggregation. *Autonomous Agents and Multi-Agents Systems*, 22(1):4–30, 2009.

N. Ailon, M. Charikar, and A. Newman. Aggregating inconsistent information: Ranking and clustering. *Journal of the ACM*, 55(5), 2008.

K. J. Arrow. *Social Choice and Individual Values.* John Wiley & Sons, second edition, 1963.

S. Barberà, H. Sonnenschein, and L. Zhou. Voting by committees. *Econometrica*, 59(3):595–609, 1991.

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.

D. Baumeister, G. Erdélyi, O. J. Erdélyi, and J. Rothe. Computational aspects of manipulation and control in judgment aggregation. In *Proceedings of the 3rd International Conference on Algorithmic Decision Theory (ADT-2013)*, 2013.

T. Bezembinder and P. van Acker. The Ostrogorski paradox and its relation to non-transitive choice. *Journal of Mathematical Sociology*, 11:131–158, 1985.

D. Black. *The Theory of Committees and Elections.* Cambridge University Press, 1958.

Borda, Jean-Charles de. *Memoire sur les Elections au Scrutin.* Histoire de l'Academie Royale des Sciences, Paris, 1781.

S. J. Brams and P. C. Fishburn. *Approval voting.* Springer, 2nd edition, 2007.

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.

S. J. Brams, D. M. Kilgour, and W. S. Zwicker. The paradox of multiple elections. *Social Choice and Welfare*, 15(2):211–236, 1998.

M. Caminada and G. Pigozzi. On judgment aggregation in abstract argumentation. *Autonomous Agents and Multi-Agent Systems*, 22(1):64–102, 2011.

F. Cariani, M. Pauly, and J. Snyder. Decision framing in judgment aggregation. *Synthese*, 163(1):1–24, 2008.

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.

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.

S. Coste-Marquis, C. Devred, S. Konieczny, M.-C. Lagasquie-Schiex, and P. Marquis. On the merging of dung's argumentation systems. *Artificial Intelligence*, 171:740–753, 2007.

T. R. Daniëls and E. Pacuit. A general approach to aggregation problems. *Journal of Logic and Computation*, 19(3):517–536, 2009.

B. A. Davey and H. A. Priestley. *Introduction to Lattices and Order*. Cambridge University Press, 2nd edition, 2002a.

B. A. Davey and H. A. Priestley. *Introduction to Lattices and Order*. Cambridge University Press, 2nd edition, 2002b.

R. Deb and D. Kelsey. On constructing a generalised Ostrogorski paradox: necessary and sufficient conditions. *Mathematical Social Sciences*, 14:161–174, 1987.

F. Dietrich. A generalised model of judgment aggregation. *Social Choice and Welfare*, 28(4):529–565, 2007.

F. Dietrich. Scoring rules for judgment aggregation. *Social Choice and Welfare*, 42(4):873–911, 2014.

F. Dietrich and C. List. Judgment aggregation by quota rules: Majority voting generalized. *Journal of Theoretical Politics*, 19(4):391–424, 2007a.

F. Dietrich and C. List. Arrow's theorem in judgment aggregation. *Social Choice and Welfare*, 29(1):19–33, 2007b.

F. Dietrich and C. List. Strategy-proof judgment aggregation. *Economics and Philosophy*, 23(3):269–300, 2007c.

F. Dietrich and P. Mongin. The premiss-based approach to judgment aggregation. *Journal of Economic Theory*, 145:562–582, 2010.

E. Dokow and R. Holzman. Aggregation of binary evaluations for truth-functional agendas. *Social Choice and Welfare*, 32(2):221–241, 2009.

E. Dokow and R. Holzman. Aggregation of binary evaluations. *Journal of Economic Theory*, 145(2):495–511, 2010.

C. Duddy and A. Piggins. A measure of distance between judgment sets. *Social Choice and Welfare*, 39(4):855–867, 2012.

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.

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.

U. Endriss and R. Fernández. Collective annotation of linguistic resources: Basic principles and a formal model. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (ACL-2013)*, August 2013. A short version was presented at BNAIC-2013 and is available at http://www.illc.uva.nl/∼ulle/pubs/files/EndrissFernandezBNAIC2013.pdf.

U. Endriss and U. Grandi. Binary aggregation by selection of the most representative voter. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI-2014)*, July 2014a. To appear.

U. Endriss and U. Grandi. Collective rationality in graph aggregation. In *Proceedings of the 21st European Conference on Artificial Intelligence (ECAI-2014)*, 2014b.

U. Endriss, U. Grandi, and D. Porello. Complexity of judgment aggregation. *Journal of Artificial Intelligence Research*, 45:481–514, 2012.

P. C. Fishburn. Arrow's impossibility theorem: Concise proof and infinite voters. *Journal of Economic Theory*, 2(1):103–106, 1970.

W. Gaertner. *Domain Conditions in Social Choice Theory*. Cambridge University Press, 2001.

W. Gaertner. *A Primer in Social Choice Theory*. Oxford University Press, 2006.

J. Geanakoplos. Three brief proofs of Arrow's impossibility theorem. *Economic Theory*, 26(1):211–215, 2005.

A. Gibbard. Social choice and the Arrow conditions. Mimeo, 1969.

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 U. Endriss. Lifting integrity constraints in binary aggregation. *Artificial Intelligence*, 199-200:45–66, 2013a.

U. Grandi and U. Endriss. First-order logic formalisation of impossibility theorems in preference aggregation. *Journal of Philosophical Logic*, 42(4):595–618, 2013b.

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. Grandi, D. Grossi, and P. Turrini. Pre-vote negotiations and binary voting with constraints. In *Proceedings of the 11th Conference on Logic and the Foundations of Game and Decision Theory (LOFT-2014)*, 2014a.

U. Grandi, A. Loreggia, F. Rossi, and V. Saraswat. From sentiment analysis to preference aggregation. In *Proceedings of the International Symposium on Artificial Intelligence and Mathematics (ISAIM-2014)*, 2014b.

D. Grossi. Unifying preference and judgment aggregation. In *Proceedings of the 8th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-2009)*, 2009.

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.

D. Grossi and G. Pigozzi. *Judgment Aggregation: A Primer*. Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool Publishers, 2014.

G. Guilbaud. Les théories de l'intérêt général et le problème logique de l'aggrégation. *Economie appliquée*, 5:501–505, 1952.

L. A. Hemachandra. The strong exponential hierarchy collapses. *Journal of Computer and System Sciences*, 39(3):299–322, 1989.

F. Herzberg and D. Eckert. Impossibility results for infinite-electorate abstract aggregation rules. *Journal of Philosophical Logic*, 41:273–286, 2012.

J. S. Kelly. The Ostrogorski paradox. *Social Choice and Welfare*, 6(1):71–76, 1989.

J. Kemeny. Mathematics without numbers. *Daedalus*, 88:577–591, 1959.

C. Kenyon-Mathieu and W. Schudy. How to rank with few errors. In *Proceedings of the 39th Annual ACM Symposium on Theory of Computing (STOC-2007)*, 2007.

A. Kirman and D. Sondermann. Arrow's theorem, many agents, and invisible dictators. *Journal of Economic Theory*, 5(2):267–277, 1972.

S. Konieczny and R. Pino Pérez. Merging information under constraints: A logical framework. *Journal of Logic and Computation*, 12(5):773–808, 2002.

L. A. Kornhauser and L. G. Sager. Unpacking the court. *Yale Law Journal*, 96 (1):82–117, 1986.

L. A. Kornhauser and L. G. Sager. The one and the many: adjudication in collegial courts. *California Law Review*, 81(1):1–59, 1993.

J. Kruger, U. Endriss, R. Fernández, and C. Qing. Axiomatic analysis of aggregation methods for collective annotation. In *Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2014)*, May 2014.

D. Lacy and E. M. S. Niou. A problem with referendums. *Journal of Theoretical Politics*, 12(1):5–31, 2000.

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.

J. Lang and M. Slavkovik. Judgment aggregation rules and voting rules. In *Proceedings of the 3rd International Conference on Algorithmic Decision Theory (ADT-2013)*, 2013.

J. Lang and M. Slavkovik. How hard is it to compute majority-preserving judgment aggregation rules? In *Proceedings of the 21sr European Conference on Artificial Intelligence (ECAI-2014)*, 2014.

J. Lang and L. Xia. Sequential composition of voting rules in multi-issue domains. *Mathematical Social Sciences*, 57:304–324, 2009.

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.

C. List and P. Pettit. Aggregating sets of judgments: An impossibility result. *Economics and Philosophy*, 18(1):89–110, 2002.

C. List and C. Puppe. Judgment aggregation: A survey. In *Handbook of Rational and Social Choice*. Oxford University Press, 2009.

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.

K. O. May. A set of independent necessary and sufficient conditions for simple majority decision. *Econometrica*, 20(4):680–684, 1952.

I. McLean and A. B. Urken. *Classics of social choice*. University of Michigan Press, 1995.

M. Miller and D. Osherson. Methods for distance-based judgment aggregation. *Social Choice and Welfare*, 32(4):575–601, 2009.

H. Moulin. *Axioms of Cooperative Decision Making.* Cambridge University Press, 1988.

J. F. Nash. The bargaining problem. *Econometrica*, 18:155–162, 1950.

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.

K. D. Nehring and C. Puppe. Abstract Arrowian aggregation. *Journal of Economic Theory*, 145(2):467–494, 2010.

H. Nurmi. Compound majority paradoxes and proportional representation. *European Journal of Political Economy*, 13(3):443–454, 1997.

H. Nurmi. *Voting paradoxes and how to deal with them.* Springer, 1999.

M. Ostrogorski. *La démocratie et l'organisation de partis politiques.* Calmann-Levy, Paris, 1902.

C. H. Papadimitriou. *Computational Complexity.* Addison Wesley, 1994.

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.

G. Pigozzi. Belief merging and the discursive dilemma. *Synthese*, 152(2):285–298, 2006.

R. A. Pollak. Bergson-Samuelson social welfare functions and the theory of social choice. *The Quarterly Journal of Economics*, 93(1):73–90, 1979.

D. Porello. Ranking judgments in Arrow's setting. *Synthese*, 173:199–210, 2010.

C. Qing, U. Endriss, R. Fernández, and J. Kruger. Empirical analysis of aggregation methods for collective annotation. In *Proceedings of the 25th International Conference on Computational Linguistics (COLING-2014)*, August 2014. To appear.

D. Rae and H. Daudt. The Ostrogorski paradox: A peculiarity of compound majority decision. *European Journal of Political research*, 4:391–398, 1976.

I. Rahwan and F. Tohmé. Collective argument evaluation as judgement aggregation. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2010)*, 2010.

W. Riker. *Liberalism against Populism.* W.H.Freeman, 1982.

K. W. S. Roberts. Interpersonal comparability and social choice theory. *The Review of Economic Studies*, 47(2):421–439, 1980a.

K. W. S. Roberts. Social choice theory: The single-profile and multi-profile approaches. *Review of Economic Studies*, 47(2):441–50, 1980b.

J. E. Roemer. *Theories of distributive justice*. Harvard University Press, 1996.

A. Rubinstein. The single profile analogues to multi profile theorems: Mathematical logic's approach. *International Economic Review*, 25(3):719–730, 1984.

D. G. Saari. *Disposing Dictators, Demystifying Voting Paradoxes*. Cambridge University Press, 2008.

P. Samuelson. Arrow's mathematical politics. In S. Hook, editor, *Human Values and Economic Policy: A Symposium*. New York University Press, 1967.

A. K. Sen. *Collective choice and social welfare*. Holden-Day, 1970.

P. Tang and F. Lin. Computer-aided proofs of Arrow's and other impossibility theorems. *Artificial Intelligence*, 173(11):1041–1053, 2009.

A. D. Taylor. *Social Choice and the Mathematics of Manipulation*. Cambridge University Press, 2005.

T. N. Tideman. Independence of clones as a criterion for voting rules. *Social Choice and Welfare*, 4(3):185–206, 1987.

N. Troquard, W. van der Hoek, and M. Wooldridge. Reasoning about social choice functions. *Journal of Philosophical Logic*, 40:473–498, 2011.

R. Vacca. Opinioni individuali e deliberazioni collettive. *Rivista internazionale di filosofia del diritto*, 52:52–59, 1922.

J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behaviour*. Princeton Univeristy Press, 1947.

K. W. Wagner. More complicated questions about maxima and minima, and some closures of NP. *Theoretical Computer Science*, 51(1-2):53–80, 1987.

R. B. Wilson. On the theory of aggregation. *Journal of Economic Theory*, 10 (1):89–99, 1975.

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.