Complexity of Winner Determination and Strategic Manipulation in Judgment Aggregation

Ulle Endriss, Umberto Grandi, and Daniele Porello

Abstract

Judgment aggregation is an area of social choice theory that analyses procedures for aggregating the judgments of a group of agents regarding a set of interdependent propositions (modelled as formulas in propositional logic). The judgment aggregation framework gives rise to a number of algorithmic problems, including (1) computing a collective judgment from a profile of individual judgments (the winner determination problem), and (2) deciding whether a given agent can influence the outcome of a judgment aggregation procedure in her favour by reporting insincere judgments (the manipulation problem). We study the computational complexity of both these problems for two concrete judgment aggregation procedures that are complete and consistent and that have been argued to be useful in practice: the premise-based procedure and (a new variant of) distance-based merging. Our results suggest that manipulating these procedures is significantly harder than solving the corresponding winner determination problem.

1 Introduction

Judgment aggregation (JA) is an area of social choice theory that analyses procedures for aggregating the judgments of a group of agents regarding a set of interdependent propositions (List and Puppe, 2009). In JA, we are given a set of propositional formulas (the agenda) and ask several agents to report which of these formulas they judge to be true. How should we aggregate this information into a collective judgment? And under what circumstances will the collective judgment be consistent? To date, most technical contributions to the JA literature have been of an axiomatic flavour, establishing characterisations and impossibility theorems (e.g. List and Pettit, 2002; Dietrich, 2006). In recent work, we have begun to investigate the computational properties of the JA framework (Endriss et al., 2010). Here, we want to extend the scope of this work and suggest a framework for analysing the computational complexity of two algorithmic problems associated with concrete JA procedures: the *winner determination problem*, i.e., the problem of computing the collective judgment from a profile of individual judgments, and the *manipulation problem*.

In the context of voting, a player is said to be able to *manipulate* a voting rule when there exists a situation in which voting in a manner that does not truthfully reflect her preferences will result in an outcome that she prefers to the outcome that would be realised if she were to vote truthfully (Gaertner, 2006). What would constitute an appropriate definition of manipulation in the context of JA? This is not immediately clear, because in JA there is no notion of preference. Here, we follow Dietrich and List (2007) and assume that a player's individual judgment set is also her most preferred outcome and amongst any two outcomes she will prefer the one that is "closer" to that most preferred outcome. We will measure "closeness" using the Hamming distance. So, we will call an aggregation procedure F *manipulable* if it permits a situation where an agent can change the outcome to a judgment set that is closer to her true judgment set by reporting untruthfully. A procedure that cannot be manipulated is called *strategy-proof*.

Dietrich and List (2007) show that F is strategy-proof if and only if it is *independent* and *monotonic*. Thus, for a meaningful study of the computational complexity of strategic manipulation, we have to restrict attention to rules that are not both independent and monotonic.[1] Furthermore, for this initial study of the subject, we choose to focus on rules that produce *consistent* and *complete* judgment sets. Specifically, we analyse two rules: the *premise-based procedure* (Kornhauser and Sager, 1993; Dietrich and Mongin, 2010) and (a new variant of) *distance-based merging* (Pigozzi, 2006). For both procedures, we compare the complexity of manipulation with the complexity of winner determination.

For the premise-based procedure, we show that manipulating the procedure is NP-hard, while winner determination is possible in polynomial time. Thus, misuse of the procedure is significantly harder than using it in the intended manner (under the common assumption that P $\neq$ NP). For distance-based merging, we show that (the decision problem corresponding to) winner determination is in NP and we conjecture that manipulation is Σ_2^p-complete (which would place the latter problem at the second level of the polynomial hierarchy). That is, under the common assumption that the polynomial hierarchy does not collapse, this would, again, make manipulation considerably harder than winner determination.

The remainder of this paper is organised as follows.[2] In Section 2 we recall the framework of JA and define the winner determination and manipulation problems. The premise-based procedure is analysed in Section 3 and distance-based merging in Section 4. We conclude with a brief discussion of related work in Section 5.

2 Judgment Aggregation

In this section we recall the basic formal framework of JA familiar from the literature (List and Pettit, 2002; Dietrich, 2006; List and Puppe, 2009) and introduce a particular notion of strategic manipulation originally proposed by Dietrich and List (2007). To make the problem amenable to a complexity-theoretic investigation, we then formulate manipulation as a decision problem, and we do the same for the winner determination problem.

2.1 The Basic Framework

We now define the basic framework for JA.[3] Let PS be a set of propositional variables, and $\mathcal{L}_{PS}$ the set of propositional formulas built from PS (using the usual connectives $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$, and the constants $\top$ and $\bot$). If α is a propositional formula, define $\sim\alpha$, the complement of α, as $\neg\alpha$ if α is not negated, and as β if $\alpha = \neg\beta$. An **agenda** is a finite nonempty set $\Phi \subseteq \mathcal{L}_{PS}$ not containing any doubly-negated formulas that is closed under complementation (i.e., if $\alpha \in \Phi$ then $\sim\alpha \in \Phi$). Denote with Φ^+ the set of positive formulas in Φ. A **judgment set** J on an agenda Φ is a subset of the agenda $J \subseteq \Phi$. Define $J(\varphi) = 1$ if $\varphi \in J$, and $J(\varphi) = 0$ if $\varphi \notin J$. We call a judgment set J **complete** if $\alpha \in J$ or $\sim\alpha \in J$ for all $\alpha \in \Phi$; **complement-free** if for no $\alpha \in \Phi$ both α and $\sim\alpha$ are in J; and **consistent** if there exists an assignment that makes all formulas in J true. Denote with $\mathcal{J}(\Phi)$ the set of all complete consistent subsets of Φ. Given a set $N = \{1, \ldots, n\}$ of $n \geqslant 3$ **agents**, denote with $\mathbf{J} = (J_1, \ldots, J_n)$ a **profile** of judgment sets, one for each agent.

Definition 1 (Aggregation procedure)**.** *A (resolute) aggregation procedure for an agenda Φ and a set of n individuals is a function $F : \mathcal{J}(\Phi)^n \rightarrow 2^\Phi$.*

[1] Independent and monotonic aggregation procedures are not very attractive: they are either dictatorial or risk producing inconsistent outcomes unless the agenda is structurally very simple (List and Puppe, 2009).

[2] We shall assume familiarity with the basics of complexity theory up to the notion of NP-completeness (see e.g. Papadimitriou, 1994). We also make reference to two complexity classes at the second level of the polynomial hierarchy: Σ_2^p, the class of problems for which a certificate can be verified in polynomial time by a machine equipped with an NP oracle, and Π_2^p, the class of problems that are complements of those in Σ_2^p.

[3] Following our earlier work (Endriss et al., 2010), to allow for a precise analysis of the computational aspects of JA, we make slight changes to the standard framework (see e.g. List and Puppe, 2009): e.g., we allow for tautologies in the agenda and we make a clear distinction between purely "syntactic" and "logical" criteria (complement-freeness *vs.* consistency). We also permit irresolute JA procedures.

That is, F maps each profile of individual judgment sets to a collective judgment set. (In Section 4 we will also introduce an **irresolute** procedure that returns a *set of* collective judgment sets.) An aggregation procedure F, defined on an agenda Φ, is said to be complete (complement-free, consistent) if $F(\mathbf{J})$ is complete (complement-free, consistent) for every $\mathbf{J} \in \mathcal{J}(\Phi)$. Here, we are only interested in procedures that are complete and consistent (and thus also complement-free). As discussed at length in the literature, these are not easy criteria to satisfy. The **majority rule**, for instance, which accepts a formula if and only if a majority of agents do, fails to satisfy consistency (Kornhauser and Sager, 1993).

Axioms provide a normative framework in which to state what the desirable (or essential) properties of aggregation procedures are. Important axioms include **anonymity**, stating that the procedure should treat all agents the same; **neutrality**, requiring symmetry with respect to propositions; **independence**, postulating that collective acceptance of φ should only depend on individual acceptance patterns of φ; and **monotonicity**, specifying that additional support for a collectively accepted formula φ should never cause φ to get rejected.[4] While all of these axioms are intuitively appealing, several *impossibility theorems*, establishing inconsistencies between certain combinations of axioms with other desiderata, have been proved in the literature. The original impossibility theorem of List and Pettit (2002), for instance, shows that there can be no consistent and complete aggregation procedure satisfying anonymity, neutrality, and independence.

2.2 Strategic Manipulation

We now define the notion of strategic manipulation for JA sketched in the introduction. Our definition is an instance of a more general definition proposed by Dietrich and List (2007), which is based on the idea that we can induce a preference relation over judgment sets by assuming that an agent's true judgment set J is her most preferred outcome, and between any two outcomes the one that is "closer" to J is preferred. One of the most appealing choices for such a notion of "closeness" is the Hamming distance.

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

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

That is, $H(J, J')$ is an integer between 0 (complete agreement) and $\frac{|\Phi|}{2}$ (complete disagreement). For example, if the agenda is $\Phi = \{p, \neg p, q, \neg q, p \wedge q, \neg(p \wedge q)\}$, then the Hamming distance between $J = \{\neg p, q, \neg(p \wedge q)\}$ and $J' = \{p, \neg q, \neg(p \wedge q)\}$ is $H(J, J') = 2$. Intuitively, if J_i is the true judgment set of agent i, then i "prefers" J over J' if $H(J_i, J) < H(J_i, J')$.

Definition 3 (Manipulability)**.** *Let Φ be an agenda, let $F : \mathcal{J}(\Phi)^n \to 2^{\Phi}$ be an aggregation procedure for that agenda, and let $\mathbf{J} = (J_1, \ldots, J_i, \ldots, J_n) \in \mathcal{J}(\Phi)^n$ be a profile. Then F is said to be manipulable at $\mathbf{J}$, if there exist an alternative judgment set $J'_i \in \mathcal{J}(\Phi)$ for some agent $i \in N$ such that $H(J_i, F(J'_i, \mathbf{J}_{-i})) < H(J_i, F(\mathbf{J}))$.*

That is, by reporting J'_i rather than her truthful judgment set J_i, agent i can achieve the outcome $F(J'_i, \mathbf{J}_{-i})$ and that outcome is closer (in terms of the Hamming distance) to her truthful (and most preferred) set J_i than the outcome $F(\mathbf{J})$ that would get realised if she were to truthfully report J_i. A procedure that is not manipulable at any profile is called *strategy-proof.* Dietrich and List (2007) have shown that a JA procedure is strategy-proof if and only if it is independent and monotonic. Thus, to study the complexity of manipulation, we have to restrict ourselves to procedures that are either not independent or not monotonic.

[4] See (List and Puppe, 2009) or (Endriss et al., 2010) for formal presentations of these axioms.

2.3 Strategic Manipulation as a Decision Problem

To study the complexity of strategic manipulation, we formulate manipulation as a decision problem. We propose the following simple definition, parametrised by the judgment aggregation procedure F under consideration.

Manipulable(F)
Instance: Agenda Φ, judgment set $J_i \in \mathcal{J}(\Phi)$, partial profile $\mathbf{J}_{-i} \in \mathcal{J}(\Phi)^{n-1}$.
Question: Is there a $J_i' \in \mathcal{J}(\Phi)$ s.t. $H(J_i, F(J_i', \mathbf{J}_{-i})) < H(J_i, F(J_i, \mathbf{J}_{-i}))$?

That is, agent i is the manipulator and her true judgment set is J_i. The other agents' judgments are given by $\mathbf{J}_{-i}$. If agent i does not manipulate, then the outcome will be $F(J_i, \mathbf{J}_{-i})$, and the Hamming distance of this outcome to her most preferred outcome (which is also J_i) is $H(J_i, F(J_i, \mathbf{J}_{-i}))$. The question we are asking is whether there exists another judgment set J_i' that agent i could report instead that would lead to an outcome $F(J_i', \mathbf{J}_{-i})$ that is closer to J_i in terms of the Hamming distance. That is, we are asking *whether* she can manipulate successfully, rather than *how*.

2.4 Winner Determination as a Decision Problem

Next, we also formulate winner determination as a decision problem:

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

By solving WinDet once for each formula in the agenda, we can compute the collective judgment set from an input profile (and, *vice versa*, any algorithm for computing the collective judgment set can be used to solve WinDet).

3 Premise-based Judgment Aggregation

There are two basic (types of) JA procedures that (can be set up so as to) produce consistent outcomes that have been discussed in the JA literature from its very beginnings, namely the *premise-based* (or *issue-based*) and the *conclusion-based* (or *case-based*) procedure (Kornhauser and Sager, 1993; Dietrich and Mongin, 2010). The basic idea is to divide the agenda into premises and conclusions. In the premise-based procedure, we apply the majority rule to the premises and then infer which conclusions to accept given the collective judgments regarding the premises;[5] under the conclusion-based procedure we directly ask the agents for their judgments on the conclusions and leave the premises unspecified in the collective judgment set. That is, the conclusion-based procedure does not result in complete outcomes, which is why we shall not consider it any further here. The premise-based procedure, on the other hand, can be set up in a way that guarantees consistent and complete outcomes, which provides a usable procedure of some practical interest—despite its well-documented shortcomings (Kornhauser and Sager, 1993; Pigozzi, 2006).

In this section, we first formally introduce the precise variant of the premise-based procedure we shall analyse. We then study the complexity of the winner determination and manipulation problems for this procedure. For ease of exposition, throughout this section, we shall assume that the number of agents n is odd.

[5] This is what is commonly understood by "premise-based procedure". Dietrich and Mongin (2010), who call this rule *premise-based majority voting*, have also investigated a more general class of premise-based procedures in which the procedure used to decide upon the premises need not be the majority rule.

3.1 Definition of the Procedure

For many JA problems, it will be natural to divide the agenda into premises and conclusions.

Definition 4 (Premise-based procedure)**.** *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. The premise-based procedure* PBP : $\mathcal{J}(\Phi)^n \to 2^\Phi$ *for Φ_p and Φ_c is the function mapping each profile $\mathbf{J} = (J_1, \ldots, J_n) \in \mathcal{J}(\Phi)^n$ to the following judgment set:*

$$\begin{aligned} \mathrm{PBP}(\mathbf{J}) \quad &= \quad \Delta \cup \{\varphi \in \Phi_c \mid \Delta \models \varphi\}, \\ & \quad\quad \textit{where } \Delta = \{\varphi \in \Phi_p \mid \#\{i \mid \varphi \in J_i\} > \frac{n}{2}\} \end{aligned}$$

If we want to ensure that the PBP always returns judgment sets that are consistent and complete, then we have to impose certain restrictions:

- If we want to guarantee *consistency*, we have to impose restrictions on the premises. It is well-known that the majority rule is guaranteed to be consistent if and only if the agenda Φ satisfies the so-called *median property*, i.e., if every inconsistent subset of Φ has itself an inconsistent subset of size $\leqslant 2$ (Nehring and Puppe, 2007; List and Puppe, 2009). This result immediately transfers to the PBP: it is consistent if and only if the set of premises satisfies the median property.

- If we want to guarantee *completeness*, we have to impose restrictions on the conclusions: for any assignment of truth values to the premises, the truth value of each conclusion has to be fully determined.

Deciding whether a set of formulas satisfies the median property is known to be Π_2^p-hard (Endriss et al., 2010). That is, in its most general form, deciding whether the PBP can be applied correctly is a highly intractable problem (and, as we shall see, a problem that is most likely considerably harder than either using or manipulating the PBP). For a meaningful analysis, we therefore restrict attention to the following case. First, we assume that the agenda Φ is closed under propositional variables: $p \in \Phi$ for any propositional variable p occurring within any of the formulas in Φ. Second, we equate the set of premises with the set of literals. Clearly, the above-mentioned conditions for consistency and completeness are satisfied under these assumptions.

So, to summarise, the procedure we consider in this section is defined as follows: Under the assumption that the agenda is closed under propositional variables, the PBP accepts a literal ℓ if and only if more individual agents accept ℓ than do accept $\sim\!\ell$, and the PBP accepts a compound formula if and only if it is entailed by the accepted literals. For consistent and complete profiles, and under the assumption that n is odd, this leads to a resolute JA procedure that is consistent and complete.

3.2 Winner Determination

Winner determination is a tractable problem for the premise-based procedure:

Proposition 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. $\square$

3.3 Strategic Manipulation

Manipulating the premise-based procedure, on the other hand, is intractable:

Theorem 2. MANIPULABILITY(PBP) *is NP-complete.*

Proof. We first establish NP-membership. An untruthful judgment set J_i' yielding a preferred outcome can serve as a certificate. Checking the validity of such a certificate means checking that (a) J_i' is actually a complete and consistent judgment set and that (b) the outcome produced by J_i' is better than the outcome produced by the truthful set J_i. As for (a), checking completeness is easy. Consistency can also be decided in polynomial time: for every propositional variable p in the agenda, J_i' must include either p or $\neg p$; this admits only a single possible model; all that remains to be done is checking that all compound formulas in J_i' are satisfied by that model. As for (b), we need to compute the outcomes for J_i and J_i' (by Proposition 1, this is polynomial), compute their Hamming distances from J_i, and compare those two distances.

Next, we prove NP-hardness by reducing SAT to MANIPULABILITY(PBP). Suppose we are given a propositional formula φ and want to check whether it is satisfiable. We will build a judgment profile for three agents such that the third agent can manipulate the aggregation if and only if φ is satisfiable. Let $p_1, \ldots, p_m$ be the propositional variables occurring in φ, and let q_1, q_2 be two additional propositional variables. Define an agenda Φ that contains all atoms $p_1, \ldots, p_m, q_1, q_2$ and their negation, as well as $m + 2$ syntactic variants of the formula $q_1 \vee (\varphi \wedge q_2)$ and their negation. For instance, if $\psi := q_1 \vee (\varphi \wedge q_2)$, we might use the syntactic variants ψ, $\psi \wedge \top$, $\psi \wedge \top \wedge \top$, and so forth. The judgment profile $\mathbf{J}$ is defined by the following table (the rightmost column has a "weight" of $m + 2$):

	p_1	p_2	$\cdots$	p_m	q_1	q_2	$q_1 \vee (\varphi \wedge q_2)$
J_1	1	1	$\cdots$	1	0	0	?
J_2	0	0	$\cdots$	0	0	1	?
J_3	1	1	$\cdots$	1	1	0	1
$F(\mathbf{J})$	1	1	$\cdots$	1	0	0	0

The judgments of agents 1 and 2 regarding $q_1 \vee (\varphi \wedge q_2)$ are irrelevant for our argument, so they are indicated as "?" in the table (but note that they can be determined in polynomial time; in particular, $J_1(q_1 \vee (\varphi \wedge q_2)) = 0$ for any φ).

If agent 3 reports her judgment set truthfully (as shown in the table), then the Hamming distance between J_3 and the collective judgment set will be $1 + (m + 2) = m + 3$. Note that agent 3 is decisive about all propositional variables (i.e., premises) except q_1 (which will certainly get rejected). Now:

- If φ is satisfiable, then agent 3 can report judgments regarding $p_1, \ldots, p_m$ that correspond to a satisfying assignment for φ. If she furthermore accepts q_2, then all $m + 2$ copies of $q_1 \vee (\varphi \wedge q_2)$ will get accepted in the collective judgment set. Thus, the Hamming distance from J_3 to this new outcome will be at most $m + 2$, i.e., agent 3 will have manipulated successfully.

- If φ is not satisfiable, then there is no way to get any of the $m + 2$ copies of $q_1 \vee (\varphi \wedge q_2)$ accepted (and q_1 will get rejected in any case). Thus, agent 3 has no means of improving over the Hamming distance of $m + 3$ she can guarantee for herself by reporting truthfully.

Hence, φ is satisfiable if and only if agent 3 can manipulate successfully, and our reduction from SAT to MANIPULABILITY(PBP) is complete. □

Thus, manipulating the premise-based procedure is significantly harder than using it—at least in terms of worst-case complexity (and under the common assumption that $\mathrm{P} \neq \mathrm{NP}$).

4 Distance-Based Judgment Aggregation

Pigozzi (2006) has shown that ideas from belief merging (Konieczny and Pino Pérez, 2002) can be imported into JA to yield practical aggregation procedures that are complete and consistent. Specifically, Pigozzi proposes a procedure that works roughly as follows: associate with each individual judgment set the model(s) satisfying that judgment set; merge the resulting set of models to obtain a new collection of models that minimise the sum of the (minimal) Hamming distances to the individual models; and return a collective judgment set corresponding to that collection of models. In this section, we introduce a new variant of this procedure and we study the computational complexity of its winner determination and manipulation problems.

4.1 A New Procedure: "Syntactic" Distance-based Merging

The merging procedure of Pigozzi (2006) has the drawback of being defined for a somewhat restricted class of profiles: the agenda is assumed to be closed under propositional variables and all compound formulas (the integrity constraints) are unanimously accepted (or rejected) by all agents. Most importantly, the syntactic information contained in the agenda is discarded by moving the aggregation from the level of formulas to the level of models. Our own proposal for distance-based merging in JA consists of a syntactic variant of this procedure, where we merge judgment sets rather than models corresponding to judgment sets. It is an *irresolute* procedure, returning a (nonempty) set of collective judgment sets.

Definition 5 (Distance-based procedure)**.** *Given an agenda* Φ*, the distance-based procedure* DBP *is the function mapping each profile* $\mathbf{J} = (J_1, \ldots, J_n) \in \mathcal{J}(\Phi)^n$ *to the following set of judgment sets:*

$$\mathrm{DBP}(\mathbf{J}) \quad = \quad \arg \min_{J \in \mathcal{J}(\Phi)} \sum_{i=1}^{n} H(J, J_i)$$

A collective judgment set under the DBP minimises the amount of disagreement with the 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). In all other profiles the consistent judgment sets that are the closest with respect to the Hamming distance are chosen as collective outcomes.

The DBP can be made *resolute* by introducing a tie-breaking rule (e.g., a lexicographic tie-breaking rule). Note that the DBP does not coincide with the procedure of Pigozzi (2006), even for agendas closed under propositional variables. The main reason is that the DBP is sensitive to logical correlations between formulas of the agenda: accepting an atom that is correlated with other formulas in the agenda "counts" more in our procedure than accepting an independent one. We find this an appealing property for a JA procedure, since it does not discard the syntactic information contained in the agenda. Also note that the DBP shares many features with the *Kemeny rule* for preference aggregation (Kemeny, 1959). We will elaborate more on this similarity in the proof of Lemma 4.

4.2 Winner Determination

Next, we want to analyse the complexity of the winner determination problem for the DBP. As the DBP is not resolute, we cannot work with the decision problem WinDet(DBP). The reason is that when there is more than one winning set, each query to WinDet (to settle the assignment for one formula at a time) may relate to a *different* winning set. We therefore formulate a new decision problem specifically for the DBP:

$\textsc{WinDet}^\star(\text{DBP})$
Instance: Agenda Φ, profile $\mathbf{J} \in \mathcal{J}(\Phi)^n$, formula $\varphi \in \Phi$, $K \in \mathbb{N}$.
Question: Is there a $J^\star \in \mathcal{J}(\Phi)$ with $\varphi \in J^\star$ s.t. $\sum_{J \in \mathbf{J}} H(J^\star, J) \leqslant K$?

That is, we ask whether there is a J with Hamming distance at most K that accepts φ. To see that this is an appropriate formulation for a decision problem corresponding to the task of computing *some* winning set, note that we can compute a winner using a polynomial number of queries to $\textsc{WinDet}^\star(\text{DBP})$ as follows. We first use it to find the smallest K for which φ_1 can be accepted, as well as the smallest K for which $\sim\varphi_1$ can be accepted ($n \cdot m$ is an obvious upper bound for K, so this can be done with a polynomial number of queries). Then we accept either φ_1 or $\sim\varphi_1$, whichever did yield the smaller K (choose either one in case of a tie). Now leave K fixed for the rest of the process. Next, substitute φ_1 with the appropriate truth value throughout $\mathbf{J}$. Then check whether φ_2 can be accepted yielding distance K; if not, $\sim\varphi_2$ must be acceptable with distance K. Accept the appropriate formula and make the appropriate substitutions in $\mathbf{J}$; then continue with φ_3, and so forth.

Unsurprisingly, the DBP is much more complex a procedure than the PBP. Nevertheless, as we show next, the complexity of winner determination does at least not exceed NP.

Lemma 3. $\textsc{WinDet}^\star(\text{DBP})$ *is in NP.*

Proof. We will show that $\textsc{WinDet}^\star(\text{DBP})$ can be modelled as an integer program (without an objective function). This proves membership in NP (Papadimitriou, 1981).

Suppose we want to answer an instance of $\textsc{WinDet}^\star(\text{DBP})$. The number of subformulas of propositions occurring in the agenda Φ is linear in the *size* (not *cardinality*) of Φ. We introduce a binary decision variable for each of these subformulas: $x_i \in \{0, 1\}$ for the ith subformula. We first write constraints that ensure that the chosen outcome will correspond to a consistent judgment set (i.e., that $J^\star \in \mathcal{J}(\Phi)$). Note that we can rewrite any formula in terms of negation, conjunction, and bi-implication without resulting in a superpolynomial (or even superlinear) increase in size.[6] So we only need to show how to encode the constraints for these connectives. The following table indicates how to write these constraints.

$\varphi_2 = \neg\varphi_1$	$x_2 = 1 - x_1$
$\varphi_3 = \varphi_1 \wedge \varphi_2$	$x_3 \leqslant x_1$ and $x_3 \leqslant x_2$ and $x_1 + x_2 \leqslant x_3 + 1$
$\varphi_3 = \varphi_1 \leftrightarrow \varphi_2$	$x_1 + x_2 \leqslant x_3 + 1$ and $x_1 + x_3 \leqslant x_2 + 1$ and $x_2 + x_3 \leqslant x_1 + 1$ and $1 \leqslant x_1 + x_2 + x_3$

Before we continue, consider the following way of rewriting the sum of distances featuring in the definition of $\textsc{WinDet}^\star(\text{DBP})$:

$$\begin{aligned}
\sum_{J \in \mathbf{J}} H(J^\star, J) &= \sum_{i=1}^{n} \sum_{\varphi \in \Phi^+} |J^\star(\varphi) - J_i(\varphi)| \\
&= \frac{1}{2} \cdot \sum_{\varphi \in \Phi} \sum_{i=1}^{n} |J^\star(\varphi) - J_i(\varphi)| \\
&= \frac{1}{2} \cdot \sum_{\varphi \in \Phi} |n \cdot J^\star(\varphi) - \sum_{i=1}^{n} J_i(\varphi)|
\end{aligned}$$

We will need to bound this sum from above. Now suppose that variables x_i with indices $i \in \{1, \ldots, m\}$ with $m = |\Phi|$ are those that correspond to the propositions that are elements of Φ. Let a_i be the number of individuals that accept the ith proposition in Φ (according

[6]For instance, any occurrence of $A \vee B$ can be rewritten as $\neg(\neg A \wedge \neg B)$. Note that rewriting a formula with nested bi-implications in terms of $\neg$ and $\wedge$ alone may result in an exponential blow-up.

to **J**). To compute a winner under the DBP, we need to find a consistent judgement set $J^\star$ (characterised by variables $x_1, \dots, x_m$) that minimises the sum $|n \cdot x_1 - a_1| + \dots + |n \cdot x_m - a_m|$. We do this by introducing an additional set of integer variables $y_i > 0$ for $i = 1, \dots, m$. We can ensure that $y_i = |n \cdot x_i - a_i|$ by adding the following constraints:

$$\begin{array}{ll} (\forall i \leqslant m) & n \cdot x_i - a_i \leqslant y_i \\ (\forall i \leqslant m) & a_i - n \cdot x_i \leqslant y_i \end{array}$$

Now the sum $\frac{1}{2} \cdot \sum_{i=1}^{m} y_i$ corresponds to the Hamming distance between the winning set and the profile. To ensure it does not exceed K, we can add the following constraint:

$$\frac{1}{2} \cdot \sum_{i=1}^{m} y_i \leqslant K$$

Finally, let $x_{i^\star}$ be the the variable corresponding to the formula $\varphi \in \Phi$ for which we want to answer WinDet*(DBP). We can force that φ gets accepted by adding one last constraint:

$$x_{i^\star} = 1$$

Now, by construction, the integer program we have presented is feasible if and only if the instance of WinDet*(DBP) we have started out with should be answered in the positive. □

Our proof also produces an algorithm for performing distance-based merging in practice. Observe that the following integer program (now *with* an objective function) can be used to find (some) winning judgment set under the DBP:

$$\min \sum_{i=1}^{m} y_i \quad \text{subject to all of the above constraints}$$

The solution can be read off from the values of the x_i. Note that the implementation details of the IP solver used will implicitly determine a tie-breaking rule. If required, other tie-breaking rules can be implemented explicitly.

Next, we show that the upper bound established by Lemma 3 is tight. Here, the similarity of the DBP to the Kemeny rule in preference aggregation allows us to build on a known NP-hardness result from the literature (Bartholdi et al., 1989b; Hemaspaandra et al., 2005).

Lemma 4. WinDet*(DBP) *is NP-hard.*

Proof sketch. We build a reduction from the problem Kemeny Score, as defined by Hemaspaandra et al. (2005). An instance of this problem consists of a set of candidates C, a profile of linear orders[7] $\mathbf{P} = (P_1, \dots, P_n)$ over C, a designated candidate c, and a positive integer K. The *Kemeny score* of candidate c is given by the following expression:

$$\mathit{KemenyScore}(c, \mathbf{P}) = \min\{ \sum_{i=1}^{n} d(P_i, Q) \mid \mathrm{top}(Q) = c \}$$

where $d(P_i, Q)$ is the Hamming distance between preference profiles and $\mathrm{top}(Q)$ is the most preferred candidate. The problem asks whether the Kemeny score of c is less than K.

We now build an instance of WinDet*(DBP) to decide this problem. Define an agenda Φ_C in the following way. First add propositional variables p_{ab} for all ordered pairs of candidates a, b in C; these variables can encode a linear order over C as a binary relation

[7] Although the Kemeny rule is defined for weak orders, the problem is known to remain NP-complete also in the case of linear orders (Bartholdi et al., 1989b, Lemma 3).

(where p_{ab} stands for $a > b$). Then add m^2 (where $m = |C|$) syntactic variants of the formula $p_{ab} \wedge p_{bc} \rightarrow p_{ac}$ for all suitable combinations of ordered pairs of candidates; these formulas encode the transitivity of the linear order encoded by the first set of variables. Finally, add an additional variable top_c. Given a preference profile $\mathbf{P}$ we can build a judgment profile $\mathbf{J_P}$ by encoding all strict orders P_i over C in a judgment set J_P over Φ_C. Due to space constraints we just show this procedure for a simple example with three candidates:

$$P = \{a > b > c\} \quad \Rightarrow \quad J_P = \{p_{ab}, p_{bc}, \neg p_{ca}, \neg top_c, \text{all transitivity constraints}\}$$

To conclude, it is sufficient to notice that $d(P, Q) = H(J_P, J_Q)$ in case P and Q share the same top candidate, otherwise the difference is 1. It is therefore sufficient to ask a query to WinDet*(DBP) using $\mathbf{J_P}$ as a profile, a suitable K' as a bound, and top_c as the fixed formula φ, to obtain an answer to the initial Kemeny Score instance with parameter K. The key step is to notice that judgment sets encoding intransitive preferences will not be considered in the minimisation process, since every disagreement on a transitivity formula will cause a much greater loss in the Hamming distance than what can be gained by modifying the variables encoding the candidate rankings. □

Putting Lemma 3 and 4 together yields a complete characterisation of the complexity of winner determination under distance-based merging:

Theorem 5. WinDet*(DBP) *is NP-complete.*

4.3 Strategic Manipulation

Next, we discuss the complexity of manipulating the DBP. Note that our definition of manipulation was tailored to resolute aggregation procedures, while the DBP (in its most general form) is irresolute and may return a set of winners. One interesting line of research to pursue in future work would be to define appropriate notions of manipulation and strategy-proofness for irresolute JA procedures. Here, instead, we shall assume that the DBP comes with a fixed tie-breaking rule (say, a lexicographic rule, or even the tie-breaking rule implicit in the IP formulation of the procedure given above, for a specific IP implementation). We do assume that this tie-breaking rule does not increase the complexity of winner determination beyond NP (this is the case for the two examples mentioned). Let Manipulability(DBP^t) be the manipulation problem for the DBP with such a fixed tie-breaking rule.

Establishing the precise complexity of manipulation for distance-based merging is currently an open problem. However, we are able to provide an upper bound:

Lemma 6. Manipulability(DBP^t) *is in* Σ_2^p.

Proof sketch. To show membership in Σ_2^p we need to show that it is possible to verify a certificate in polynomial time on a machine that has access to an NP oracle. Recall from the first part of the proof of Theorem 2 that an appropriate certificate is a judgment set J_i' for the manipulator that is complete and consistent and that produces an outcome that is closer to the manipulator's true judgment set J_i than the outcome produced if she reports J_i. This involves three non-trivial steps, all of which can be resolved by the NP oracle: deciding consistency of J_i' is in NP (this is just Sat), and computing the winners for J_i and J_i' is also in NP (by Lemma 3). Thus, the certificate can be verified using three calls to the oracle; the remainder of the computation is clearly polynomial. □

We conjecture that the above bound is tight, i.e., that Manipulability(DBP^t) is Σ_2^p-complete.[8] If this conjecture is correct, then manipulation is significantly harder than winner determination, also in the case of distance-based merging.

[8] To the best of our knowledge, there are currently no known results on the complexity of the (presumably) closely related problem of manipulating Kemeny elections.

5 Related Work

We conclude by briefly reviewing some related work regarding (1) alternative notions of manipulation in JA, (2) other complexity-theoretic questions in JA, (3) manipulation and strategy-proofness in belief merging, and (4) the complexity of manipulation in voting.

As mentioned earlier, our definition of strategic manipulation in JA is based on the work of Dietrich and List (2007). This definition crucially rests on the idea that we can induce a preference ordering over judgment sets from an agent's true judgment set an a metric for measuring "closeness". The Hamming distance is one such metric; Dietrich and List (2007) also discuss the concept of "closeness-respecting" preferences (and the corresponding notions of strategic manipulation) in more general terms. Other than that there has been preciously little work on manipulation in JA to date. One exception is the work of Pigozzi et al. (2009), who introduce a notion of *full manipulability*, which asks whether an agent can change the outcome to fully coincide with her own judgment set by means of an insincere judgment. But (as clearly recognised by the authors) the guarantee of the absence of full manipulation is probably a property that is simply too easy to satisfy to lead to interesting characterisations of JA procedures.

In previous work (Endriss et al., 2010), we have analysed the complexity of another aspect of the JA framework: for a given set of axioms characterising a class of aggregation procedures, how hard is it to check whether a given agenda is *safe* for all procedures belonging to that class, in the sense that no profile of complete and consistent individual judgment sets will ever result in a collective judgment set that is not consistent? (Our results suggest that deciding *safety of the agenda* is Π_2^p-complete for most natural combinations of the standard axioms.) To the best of our knowledge, this is the only other work on the computational complexity of JA to date.

The field of *belief merging* is closely related to judgment aggregation (Konieczny and Pino Pérez, 2002; Pigozzi, 2006). A definition of strategy-proofness for belief merging operators has been proposed by Everaere et al. (2007), and the same authors have discussed the problem of manipulation for a range of belief merging operators. While this work does include the study of the complexity of belief merging, the complexity of manipulation has, to the best of our knowledge, not yet been addressed in the belief merging literature.

Finally, there are of course close connections between our work and the line of work in computational social choice that has studied the complexity of both the winner determination and the manipulation problem for a range of voting rules in depth, starting with the seminal work of Bartholdi et al. (1989a,b). Some of this work has been reviewed by Chevaleyre et al. (2007), who give many references. Recent discussion in the literature on the complexity of manipulation of elections has centred on the question of whether worst-case results (such as NP-hardness results) are sufficient deterrents against manipulation in practice (see e.g. Procaccia and Rosenschein, 2007). They probably are not; what is really needed is a better understanding of the average-case complexity of manipulation. The very same questions will have to be asked for JA as well; our (worst-case intractability) result and conjecture are only the first step.

References

J. J. Bartholdi, C. A. Tovey, and M. A. Trick. The computational difficulty of manipulating an election. *Social Choice and Welfare*, 6(3):227–241, 1989a.

J. J. Bartholdi, C. A. Tovey, and M. A. Trick. Voting schemes for which it can be difficult to tell who won the election. *Social Choice and Welfare*, 6(2):157–165, 1989b.

Y. Chevaleyre, U. Endriss, J. Lang, and N. Maudet. A short introduction to computational social choice. In *Proc. 33rd Conference on Current Trends in Theory and Practice of Computer Science (SOFSEM-2007)*. Springer-Verlag, 2007.

F. Dietrich. Judgment aggregation: (im)possibility theorems. *Journal of Economic Theory*, 126(1):286–298, 2006.

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

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

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

P. Everaere, S. Konieczny, and P. Marquis. The strategy-proofness landscape of merging. *Journal of Artificial Intelligence Research*, 28(1):49–105, 2007.

W. Gaertner. *A Primer in Social Choice Theory.* LSE Perspectives in Economic Analysis. Oxford University Press, 2006.

E. Hemaspaandra, H. Spakowski, and J. Vogel. The complexity of Kemeny elections. *Theoretical Computer Science*, 349(3):382–391, 2005.

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

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. The one and the many: Adjudication in collegial courts. *California Law Review*, 81(1):1–59, 1993.

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.

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.

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

C. H. Papadimitriou. On the complexity of integer programming. *Journal of the ACM*, 28 (4):765–768, 1981.

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

G. Pigozzi, M. Slavkovik, and L. van der Torre. A complete conclusion-based procedure for judgment aggregation. In *Proc. 1st International Conference on Algorithmic Decision Theory (ADT-2009)*. Springer-Verlag, 2009.

A. D. Procaccia and J. S. Rosenschein. Junta distributions and the average-case complexity of manipulating elections. *Journal of Artificial Intelligence Research*, 28:157–181, 2007.

Ulle Endriss, Umberto Grandi and Daniele Porello
Institute for Logic, Language and Computation
University of Amsterdam
Postbus 94242, 1090 GE Amsterdam, The Netherlands
Email: `{u.endriss,u.grandi,d.porello}@uva.nl`