UNIVERSITÉ DU LUXEMBOURG

PhD-FSTC-2012-05
The Faculty of Sciences, Technology and Communication

DISSERTATION

Defense held on 31/01/2012 in Luxembourg

to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG
EN INFORMATIQUE

by

Marija SLAVKOVIKJ

Born on 29 December 1979 in Skopje (Macedonia)

JUDGMENT AGGREGATION FOR MULTIAGENT SYSTEMS

Dissertation defense committee

Prof. Dr. Sjouke Mauw, *Chairman*
Professor; Université du Luxembourg

Prof. Dr. Leendert van der Torre, *Dissertation Supervisor*
Professor; Université du Luxembourg

Dr. Gabriella Pigozzi
Assistant Professor; Dauphine Université Paris

Dr. Ulle Endriss
Associate Professor; University of Amsterdam

Dr. Sébastien Konieczny
Directeur de Recherche CNRS ; Centre de Recherche en Informatique de Lens

Printing supported by

Fonds National de la
Recherche Luxembourg

Cover design

Yann Flammang

Acknowledgments

I do not particularly like Charles Dickens but when describing the experience of making a doctoral thesis I must use the entry line of his *A Tale of Two Cities*: “It was the best of times, it was the worst of times”. I have heard that theses’ are never finished, rather abandoned. As I abandon mine I would like to thank the people that made working on my thesis the best of times.

I did this thesis under the supervision of Leon van der Torre and Gabriella Pigozzi. It is a great responsibility to supervise a doctoral student and I have the privilege of being the first student of Gabriela’s. I am sure it was not always easy for her but from my view point she did a great job and for this I am gratefull. I am grateful to Leon for his patience and guidance. Somehow I am sure he thought me a lot more than he intended. He thought me not only research but also management. Thanks boss!

I had the privilege of being externally supervised by Dov Gabbay. He is a legend and I am happy he was onboard my thesis. Thanks for being there!

I learned a lot from my coauthors. I would like to thank Guido Boella for being my host and teacher. I would also like to thank Wojtek Jamroga for working with me and showing me trough complexities. You two took good care of me and I would like to acknowledge that. I have had a lot of fun working with Jérôme Lang. I like his academic style I do hope I would be able to emulate it. I do hope that I would continue writing papers with all of you.

I have had the privilege of being part of a nice group that expanded and contracted over the four years I spent in Luxembourg. I have shared an office with four guys. First was Davide Grossi, the most algorithmic researcher I know. I am fascinated by his work approach. I do hope I will get to be on his level one day. Davide, I am sorry I annoyed you, but that was only in admiration. Second was Matthijs de Boer, the most non algorithmic researcher I know. A very smart guy who I enjoyed sharing an office with. Third was Valerio Genovese who as Davide is a seriously algorithmic fellow, but with whom I did not work together unfortunately. Yet. Lovely sharing an office with him. Lastly and too briefly Francois Francesco Schwarzentruber, a very nice guy and I am sorry we did not do the office share longer.

A person I did not share an office with but who had indirectly influenced my thesis is Martin Caminada and I hope our paths cross again. I would like to thank the rest of my colleagues from my group and out of my group for all the lovely moments we shared: Emil, Yining, Patrice, Richard, Xavier, Antonis, Srdzjan, Silvano, Patricio, Mikolaj, Raymond, Ton, Hugo, Sasha, Tjitze, Pouyan, Matthijs, Simon, Viji, Sergio, Loic and last but by far not least Paolo Turrini.

I would like to thank the defense committee members, Ulle, Sébastien and Sjouke, for doing a great job. They were helpful by being constructively critical when necessary. I hope I do not do them a bad favor by emphasizing that they are a great choice for this role.

The people I have so far mentioned I consider friends, in addition to being mentors and colleagues. In addition I would like to thank my out of academia friends for still being my friends when I fell off the edge of the Earth while writing this thesis. I would like to specially thank Ronan for loving me, keeping me sane and introducing me to many fun elements.

At the end I would like to thank my parents and brother for standing by my choices even when not agreeing with them.

Contents

List of Figures

List of Tables

1

Introduction

1.1 Background

Artificial intelligence is concerned with the task of enabling machines to solve complex problems by, *e.g.*, learning, behaving intelligently, reasoning, decision-making. In the seventies of the last century, when artificial intelligence was at its beginnings, these cognitive processes were seen, studied and modeled separately. Later on, results in various artificial intelligence disciplines accumulated and computer technology rapidly advanced. As a consequence, the paradigm of the *intelligent agent* became an appealing approach to study and recreate the human mental activities.

An agent is an autonomous entity that embodies several cognitive processes, defined as "anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators" (Russell and Norvig, 2010, pg.34). An intelligent agent is defined as "a computer system that is capable of independent action on behalf of its user or owner" (Wooldridge, 2009, pg. 5). By definition an agent interacts with its environment and other agents. The appeal of the agent paradigm lies precisely in the possible computational power that emerges from the interaction among agents. The structures formed by intelligent agents are called *multiagent systems*. The interest in multiagent systems particularly took off with the advent of social software when the role of a computer shifted from the computer being a self contained machine for executing software, a "personal computer", to being a "net-book", a global communication tool and an access node for disseminating information, conducting commerce and efficient leaking of embarrassing personal information to potential employers.

The interactions within a multiagent system include cooperation and coordination. To be able to coordinate and cooperate, intelligent agents need to reach collective consents, namely binding group decisions, over issues such as beliefs, actions and desires. One type of collective consent is *an agreement*. An agreement is a mutual and enforceable understanding among agents. The processes and mechanisms implicated in reaching agreements among agents have recently become a subject of research and analysis from technology-oriented perspectives (Ossowski, 2008).

The interactions among people and how they reach collective consents are studied within the scope of economic theory, by social choice theory. In economic, a decision is a choice of option(s) from a given set of options. The set of options is also sometimes called a set of alternatives. Social choice theory includes voting theory, preference aggregation and judgment aggregation. These theories are all concerned with developing and studying methods for making collective decisions.

Preference aggregation (Arrow et al., 2002, Part 1) studies the problems of forming a group opinion for a set of options. Each agent specifies which options he most prefers, which he prefers less and so on, building a subjective preference order over the set of options. A preference aggregation rule fuses these subjective orders into a preference order that is representative for the group.

Voting theory (Arrow et al., 2002, Chapter 4), (Nurmi, 2010) studies the problems of making a group choice from a set of candidates. Each agent casts a vote for or against one, some or all of the candidates. The structure of the vote depends on the voting context. The simplest vote structure is the one-person-one-vote, when each agent is allowed to choose one candidate from the candidate set. The most elaborate vote structure is a total preference order, as in preference aggregation. Voting occurs in many formal contexts such as: political elections, electing best entries in contests and determining the winners in sport competitions like figure skating. Voting also occurs in informal contexts, such as groups of people deciding where to go for dinner, how to name their robots, *etc*. A voting rule selects a winner from the set of candidates based on the individual votes.

Judgment aggregation (List and Puppe, 2009) studies the problems of making group decisions regarding the truth-value of several issues considered concurrently. For one set of issues, all combinations of truth-value assignments are not allowed. Judgment aggregation problems occur in committee and jury decision-making contexts. As in voting theory, the contexts of judgment aggregation problems range from entirely formal to entirely informal. An example of a formal context is a collegiate court which is deciding whether a given case is within the jurisdiction of a given court, whether the presented evidence for the case are sufficient for a trial, and whether a trial should be scheduled. A trial can be scheduled if and only if the evidence is sufficient and a court has jurisdiction. An example of an informal judgment aggregation context is a group of friends deciding on whether to go to a certain restaurant, whether the restaurant in question has vegetarian dishes on the menu and whether the prices are affordable. The group can only go to the restaurant if it is the group's opinion that there are vegetarian dishes and that the prices are affordable.

Each agent forms a judgment regarding the truth state of each issue. Usually a judgment is a binary value denoting whether an issue is true or false, accepted or rejected. A judgment aggregation rule aggregates these truth-value assignments into an allowed combination of truth-value assignments, one for each considered issue.

The Figure 1.1 is an abstract simplified illustration of preference aggregation (left hand-side funnel), voting (center funnel) and judgment aggregation (right hand-side funnel), and allows us to make a comparison between the three. In each of the social choice problems presented on Figure 1.1, there are three agents: Top, Middle and Bottom. Their individual preferences, votes and judgments are represented in the corresponding order. The group decisions are represented in the exit of the funnel. In the case of preference aggregation the set of options are a star, a circle and a square. In voting, the star, the circle and the square are candidates. In this picture we give the most complicated vote construct, the full preference order. In the case of judgment aggregation, the star, the square and the circle are the issues on which judgments are cast. Each agent assigns a value true (yes) or false (no) to each issue. The relations between the star, the circle and he square are such that if an agent accept either the star or the circle, then he has to accept the square as well.

Preference aggregation, voting and judgment aggregation all appear to be simple. However, these procedures, and related theories, are for many reasons, far from simple. The variety of

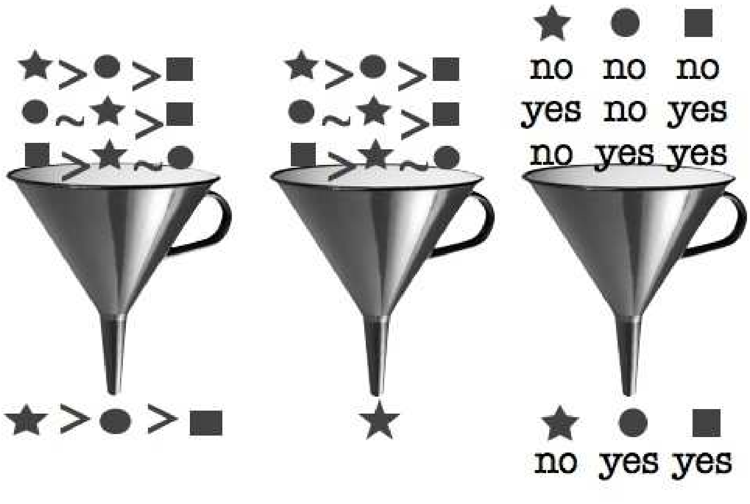


Figure 1.1: Different social choice problems: preference aggregation (left), voting (middle) and judgment aggregation (right).

contexts in which these social choice rules are used is extremely large and so is the variety of requirements that the group decision should satisfy with respect to the individual opinions, votes or judgments correspondingly. There are many different procedures that can be used for the same problem each leading to a different group decision. There also are many combinations of requirements that no procedure can satisfy simultaneously. There are problems for which some procedures can be applied more efficiently than others.

Economic theory is no stranger to computer science when it comes to applying methodology from one to the other. While social choice rules generate collectively binding group decisions, another discipline in economy, decision-making, considers the problem of making individual decisions. Chapters 16 and 17 in (Russell and Norvig, 2010) illustrate that decision theory is a staple methodology used in artificial intelligence. In the last five years, the exchange of ideas and methodologies between economics and computer science is flowing both ways, as witnessed by the very fruitful field of *algorithmic game theory* (Roughgarden, 2010).

Importing concepts from social choice theory into computing and applying computational analysis in social choice is studied by *computational social choice* (Chevaleyre et al., 2007). The first direction of using methods from computing to study problems in social choice theory is well explored. The most typical problem studied in computational social choice is the complexity-theoretic analysis of voting protocols, such as (Bartholdi et al., 1989; Hemaspaandra et al., 1997; Conitzer and Sandholm, 2002b,a; Walsh, 2008). Other typical problems include allocation of resources (Maudet, 2010, Chapter 3), (Chevaleyre et al., 2005); formal specification and verification of social procedures using mathematical logic, such as (Bouveret and Lang, 2005), (Maudet, 2010, Chapter 2); compact representation of elicited input using logic, such as (Bienvenu et al., 2010) and the computer aided search for properties of social choice rules, such as (Tang and Lin, 2009).

The second direction of importing concepts from social choice to computer science is also outlined as a core consideration of computational social choice. This direction is considerably less explored. One would expect that some of the first models for obtaining collective consents in multiagent systems would be inspired by, or imported from, social choice theory. This is not what happens.

Why is there a gap where there should be work that explores the use of social choice for collective reasoning problems in multiagent systems? Are there no multiagent systems problems that can benefit from social choice methods? We give a few examples to illustrate that this is not the case.

Argumentation based negotiation methods (Beer et al., 1999) are seen as essential in enabling agents to reach agreements that respect the constraints imposed by norms and organizations (Ossowski, 2008). However negotiation is insufficient to cover all collective reasoning problems. The most notable difference between negotiation and aggregation is in the number of information exchanges between the agents before group consent is obtained. Negotiation procedures presume a potentially unspecified number of exchanges. Aggregation requires that the agents submit their preferences, choices or judgments; correspondingly, only once to an agent or service which aggregates them.

There are contexts in which the agents cannot or will not commit to numerous exchanges of opinions. Agent teams that operate in uncertain environments, such as robots conducting rescue missions, cannot afford the time to negotiate about what to do since their options can change while they are still negotiating on which option to choose. A hierarchical group is a group in which there is one agent who responsible for making group decisions. This agent often needs to consider the opinions of other group members to reach that decision. In hierarchical groups, aggregation is sometimes a better approach than negotiation since the agents that do not make the decision can be unwilling to participate in negotiation.

A special case of a hierarchical group is an agent that needs to acquire information about the environment, by considering the opinions on other agents. Consider a robot that does not have a microphone. It needs to determine whether an alarm is on in a building and whether the alarm being on implies the need to vacate the premises. Different other robots may report different information on these two counts, or even have different opinions on whether the building has to be vacated. Our robot can aggregate the received information to determine what to do and what to believe.

The improvement of information and communication technological systems (ICT systems) depends on the evaluation of the users. The users provide feedback that is used to modify certain system's features, such as for example resilience and dependability. The feedback of the user can be different regarding the same feature. The software engineers need to analyze the user data and determine which features to modify and in which direction to modify them. The user feedback is a valuable commodity. A lot of effort has been spent on the technical support of eliciting opinions, *i.e.*, voting. Technical means are used to resolve issues such as guarantee of privacy, eliminating possibilities for coercion and security. However, once the information is obtained, engineers cannot expect that the users will negotiate it among each other and agree on which features they like to see improved. The users together with the producers form a hierarchical group. It is difficult to derive a collective consent from a multiple feedback without a formal and automatized method. Such methods are needed even when a standardized input on a set of qualifiers is used to elicit the information because certain features depend on others.

Prediction markets, also known as "event futures" and "information markets" are markets in which agents trade contracts with payoff that depend on unknown future events. The goal of designing prediction markets is to make accurate forecasts. This is done by aligning the experts' incentives with the elicitation of information and by aggregating their opinions. The agents that make predictions are myopic, have fixed beliefs about the value of a contract, and are risk neutral. They have a fixed, finite budget and participate exactly once, by acting in the market and then exiting, see for instance (Othman and Sandholm, 2010). When his prediction is confirmed, an agent is rewarded with increased weight on his prediction, and punished with reduced weight when his predictions are wrong. At each step the forecast of the agents on a set of market prices needs to be aggregated. This process is sequential and tied to real world events. The agents are presumed to be selfish, so negotiation is not an option.

In negotiation, the produced consent depends not only on the information the individual agents have but also on the negotiation skill of particular agents. When the group is heterogeneous, the input from "weaker" agents will be marginalized. Consensus groups are groups in which there is no one agent responsible for making the decision. An example of a decision-making in a consensual group is the establishing of group mental attitudes, such as beliefs and intentions. In multiagent systems, it is usually taken that a group has an attitude if every member of the group individually has the same attitude, but his is not the only way to model collective attitudes (Dunin-Keplicz and Verbrugge, 2010, Chapter 3). How collective attitudes are formed is studied in social epistemology. Social epistemology offers an alternative definition of collective attitudes: a group has an attitude if the group members agree to have that attitude, see for instance (Gilbert, 2009). This is called the *non-summative approach* to collective attitudes and it is more flexible in allowing groups to act together, as they do not need to be equally minded to have joint attitudes.

The aim of this thesis is to explore the possibilities of using social choice procedures as method for reaching collectively binding decisions in multiagent systems.

1.2 Research Question

The research question pursued in this thesis is the following:

> How can judgment aggregation operators be designed and selected for use in multi-agent systems?

The social choice rule used to combine individual opinions, votes or judgments correspondingly, can be seen as a type of a norm. The social choice rule is established before the opinions, judgments or votes, are elicited. This is necessary, since one can design a rule that produces a desired outcome from an individual input. For instance, in presidential elections, a parliament or other authority before the elections sets the rule according to which the president is elected. A rule is chosen to best serve the purposes of the context in which it is applied. This is why the challenge in using social choice to obtain group consents automatically is in the selection of adequate rules for multiagent contexts.

We enumerated three social choice disciplines developing and studying methods for generating collectively binding decisions: voting, preference aggregation and judgment aggregation. Why focus on judgment aggregation? Voting and preference aggregation are very similar, they both aggregate agents' preferences over a set of options. Voting rules produce an option that is the most preferred, *i.e.,* a winner or alternatively a set of winners, while preference aggregation rules produce a collective preference order over the set of options. The problem

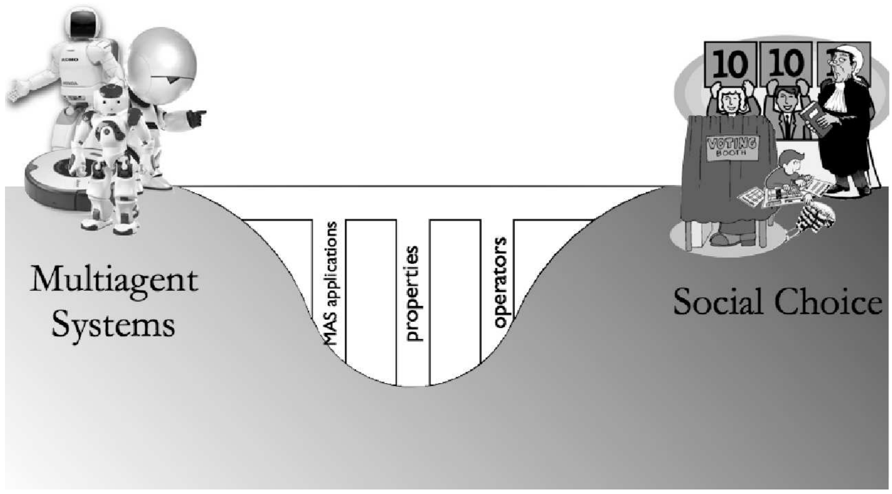


Figure 1.2: Bringing social choice theory with multi-agent systems closer.

of aggregating judgments only started attracting considerable attention in the last ten years, since (List and Pettit, 2002; List and Puppe, 2004) showed that judgment aggregation is general in the sense that it subsumes voting theory and preference aggregation (List and Polak, 2010). The novelty and the generality of judgment aggregation is the reason we have chosen it among the aggregation theories of social choice.

The thesis research question is tackled by considering three sub-problems:

1. Designing judgment aggregation rules.
2. Classifying judgment aggregation rules by properties they satisfy.
3. Pairing aggregation contexts with adequate rules.

Figure 1.2 illustrates symbolically the roles that these sub-problems have in answering the research question. We design operators for aggregating judgments and we use them in group decision problems that occur in multiagent systems. The properties of the operators and the properties of the group decision problems are used to pair one with the other.

Voting has been formally studied since the seminal works of Borda and Condorcet in the eighteenth century. Judgment aggregation is resent when compared to voting. The problem of aggregating judgments was observed by (Kornhauser and Sager, 1986), but it has its precursors in the works of (Gilbaud, 1966; Wilson, 1975) and (Rubinstein and Fishburn, 1986), see (List and Polak, 2010) for a detailed historical overview and comparison. The interest in voting theory is caused by the need to conduct democratic elections. Therefore, it was of interest to develop dozens of specific voting rules over the years. The interest in judgment aggregation was sparked when it was shown that it applies to problems that are different than the ones studied in voting theory. The majority of the work in judgment aggregation is devoted to studying impossibility results in the style of the work in preference aggregation by Arrow, see (Arrow, 1963) and (List and Puppe, 2009). A small number of specific rules for aggregating judgments have been proposed, however the development of specific aggregation rules is still widely unexplored.

It is different whether a social choice method, for instance voting, is applied in a democratic election or for the purpose of reaching automated consent. The difference is in the frequency of the process and the impact it has on the agents that use the consent produced by it. The voting in elections occurs infrequently, but the impact of the results is enduring. In general, people only vote for issues that are critical. In computational and multiagent contexts the situation is reversed. The need for consent is frequent, but the impact of the consent is low. If a group observes it is doing something wrong, it can re-vote. Eliciting information from an artificial agent is much more feasible than organizing national elections. This is why desirable properties for rules used in human social choice contexts are not necessarily desirable for the rules used in multiagent systems. For instance, the incentive of an agent to manipulate the aggregation process to obtain a desirable outcome is a big issue in elections. However, when the impact of the consent is low, then it can be expected that the incentives to manipulate are also low. There are many properties of rules studied in voting theory that ensure that the consent is desirable with respect to the individual input. Some of these properties are also desirable for judgment aggregation rules and need to be defined in terms of judgment aggregation rules. This is the core of the second sub-problem.

To recommend a judgment aggregation rule for a multiagent systems problem one needs to pair the properties of the rule with the characteristics of the problem. There are many contexts that can be specified. One general approach is to look into the type of group that needs to use the decision-making procedure. In the broadest sense of who does the collectively binding decisions apply to, we can distinguish between two types of groups: a hierarchical group and a consensual group. In hierarchical groups there is one agent that is responsible for making the decision for the group by considering the opinions of the group members. In consensual groups, no one responsible agent exists. The group decision emerges or is proposed by any of the agents. In the case of hierarchical groups, the decision can apply only to the agent who is responsible to make it, only to the agents who contribute the opinions, or to all agents within some institution. In consensual groups, all the decision makers are the decision "targets". We consider consent reaching contexts in hierarchical and consensual groups and we give one example of group decision-making procedures based on judgment aggregation for each of these contexts.

Within the scope of this theses, we consider only the case when the agents involved do not behave strategically, in both the hierarchical and consensual context. Therefore here we do not define nor study vulnerability to manipulation of the introduced rules or consent reaching models. Both the examples we give involve groups that cooperate in pursuing a group goal. This implies that the groups, in addition to reaching collectively binding decisions, also need to form joint plans and to communicate with each other. We do not consider specifics of planning and inter-agent communication. We assume that these activities are possible and do not hamper the judgment aggregation based consent reaching procedures we propose. We also do not consider group learning, although we acknowledge its relevance, particularly in-group adaptation, which we do consider. In the context of hierarchical groups, the agents are bound to acknowledge, conform and act according to the group decisions by the context in which the decision is made. For instance, all members of an institution are bound by the decisions of the president of the institution. In the context of consensual groups, the fact that the group decision is binding needs to be additionally expressed. This is the reason why we also design and study group commitment strategies in our example.

Let us elaborate each of the sub-problems, and how go about solving them, in more detail.

1.2.1 Designing judgment aggregation rules

In this section we consider the problem of designing judgment aggregation rules. A judgment aggregation problem is specified by an *agenda* and a set of agents. An agenda is a set of logically related issues, usually referred to as *propositions*. It is common to represent the agenda issues and the relations that hold between them in propositional logic. Each agent expresses an acceptance or rejection regarding individual propositions in the agenda. The expressed acceptance or rejection of an issue is called *a judgment*.

An intuitive approach to aggregating the judgments is to consider how many agents support each truth-value for each of the issues and adhere to the will of the majority. We consider the concept of majority and the different ways it can be used to construct judgment aggregation rules.

The concept of majority in judgment aggregation

Example 1.2.1. *Consider four agents* $\{Red, Blue, Green, Orange\}$ *and an agenda of four issues: a, b, c and d. The relations are such that d can be accepted if and only if a and either b or c are accepted, namely* $(a \wedge (b \vee c)) \leftrightarrow d$. *Table 1.1 represents a possible judgment aggregation problem. The* $+$ *entry denotes an accepted proposition and the* $-$ *entry, a rejected one. The collection of all judgments received from the agents is called* a profile, *which is the white panel in Table 1.1.*

Agents	**Agenda**			
	a	*b*	*c*	*d*
Red	-	+	+	-
Blue	+	-	-	-
Green	+	+	-	+
Orange	+	-	-	-

Table 1.1: An example of a judgment aggregation problem.

It is common to require completeness, namely that each agent either expresses an acceptance or rejection for each issue. Each agent is constrained by the logical relations, in the sense that the combination of issues he accepts or rejects must satisfy these constraints.

Each of the judgment sets of the agents in Table 1.1 conforms to the logic relations between the issues. For instance, Red rejects a and also rejects d, while accepting both b and c. If Red were to accept d in addition to accepting b and c, while rejecting a, his judgment set would have been inconsistent.

The first problem is to determine how to define majority in judgment aggregation. Sets of judgments are particular types of information. On one hand a set is a unit of information since it represents the opinions of one agent on one agenda. On the other hand, the set contains judgments that can be considered to be units of information.

Let us consider the judgment set as an atomic information unit. In this case, if there is a judgment set that is supported by more agents than any other judgment set, then this is a majoritarian judgment set. Let us call *set-majoritarian* the set of judgment that is selected, as a whole set, by the largest number of agents, with respect to the profile. Consider Table 1.1 as an example. Blue and Orange both accept the same set, while the sets of Red and Green are different from the Blue-Orange one and from each other. Consequently the Blue-Orange set is the set-majoritarian set. The set-majoritarian judgment set does not always exist.

Let us consider the judgments to be an atomic unit of information. In this case, if there is a judgment set in which each judgment is supported by a strict majority of agents, then this set can be considered a majoritarian judgment set. Let us call *issue-majoritarian* the set of judgments that is determined by counting the majority judgments on each issue. In which order should this issue-by-issue aggregation be done? Consider Table 1.1 as an example. If applied to all issues at once the result of this exercise is an acceptance of a, a rejection of c and d and no decision on b. If b is accepted then the collective set violates the constrains, and a majority does not exist. If b is rejected then the set is consistent with the issue relations, but there is no reason for b to be rejected? The problem is more general: due to the logic relations among the issues, for every rule that aggregates the judgments issue-by-issue can produce, there exists some profile for which the rule produces a judgment set that violates the constraints.

Some agendas can be conceptually partitioned to a set of premises and a set of conclusions. A conclusion is typically an issue whose acceptance can be deduced from the acceptances and rejections of the premises. For example, let us interpret the agenda issues as follows:

- a a victim is trapped in a location that is difficult to access
- b the victim is conscious
- c the victim is in a face-up position
- d save the victim using a rescue harness

The issues a, b and c are premises. In this case they are the necessary and sufficient conditions under which certain action d will be taken. A *premise-based procedure* is the aggregation rule that calculates the majority for each premise and deduces whether the conclusion is accepted or rejected based on the issue relations.

The premise-based procedure is an appealing alternative to the issue-by-issue aggregation, however there are some problems with using it. The problem with the premise-based procedure is that:
a) not every agenda can be conceptually partitioned intro premises and conclusions, and
b) even if the partitioning is possible, the conclusion is not deducible from the premises in every set of judgment sets.

Consider the profile in Table 1.1. Using the premise-based aggregation rule we obtain that a

is accepted and c rejected. We get no decision for b and cannot deduce the decision for the conclusion d.

Using majority and minimization to design rules

When a consensual group needs to reach decisions, these decisions should reflect the "will of the majority" for them to be acceptable to the group. Therefore, one would like to have rules that select the issue-majoritarian or the set-majoritarian judgment. The problem is that neither the set-majoritarian nor the issue-majoritarian sets exist for every profile, but every set-majoritarian judgment set is an issue-majoritarian judgment set and the reverse does not hold. We can have the second-best thing: a rule that selects the issue-majoritarian judgment set whenever such a set is consistent with the constraints.

The appeal of the majoritarian sets is that, when selected as collective consent, they *minimize* the discrepancy between the collective consent and the elicited information. It is this minimal discrepancy that is desirable for consensual groups and we use it to design judgment aggregation rules. We call *majority-consistent* any profile for which an issue-majoritarian judgment set exists. What if we change the profile in some minimal way so that it becomes majority-consistent and then select its issue-majoritarian judgment set as the collective judgment set for the original profile? There are many ways to minimally alter a profile. For example, this can be done by removing judgments on an issue, individual judgments, judgments that belonging to some agent(s), by repeating judgments etc. There can be, as many rules as there are minimal alterations to a profile that can be defined, but each of these rules will by construction always select the issue-majoritarian set when such a set exists for the starting, unaltered, profile.

There is a third way to use majority, we can treat the judgment sets as units that are qualified by the individual judgments they consist of. This allows us to define a measure of similarity, or *distance*, between two judgment sets that is finer grained than [equal, different]. The majority concept here corresponds to most similar. What does it mean that a judgment is most similar to a profile of judgments?

The similarity between sets can be quantified based on the number and the type of judgments on which the two sets differ. For instance, the sets of Red and Blue differ on three judgments, and so do the sets of Red and Green. However Red and Green differ on different three judgments. Red and Blue give the same judgment for what is the conclusion under the interpretation we gave. Hence, the set of Red can be considered more similar to the set of Blue than to the set of Green. Given the context of the aggregation problem many similarity measures can be specified.

Numeric distances can be aggregated by using an arithmetic aggregation function. The collective consent is the set that is closest to all the individual judgment sets. The interpretation of how close is a judgment set to a profile of judgments is set by the arithmetic aggregation function. Table 1.2 illustrates an aggregation of distances using the sum as an arithmetic aggregator and the number of different judgments as a similarity measure between two judgment sets.

We consider not only the distance from each of the contributed sets to the profile, but also the distances from any judgment set that satisfies the issue relations to the profile. In the left handside of Table 1.2, all the acceptable judgment sets, for the agenda and constraints in Example 1.2.1, are enlisted. The numbers under each agent's name, the number indicates on how many

a	*b*	*c*	*d*	Red	Blue	Green	Orange	∑
-	-	-	-	2	1	3	1	7
-	-	+	-	1	2	4	2	9
-	+	-	-	1	2	2	2	7
-	+	+	-	0	3	3	3	9
+	-	-	-	3	0	2	0	5
+	-	+	+	3	2	2	2	9
+	+	-	+	3	2	0	2	7
+	+	+	+	2	3	1	3	9

Table 1.2: Quantifying the similarities between judgment sets and aggregating them using the arithmetic aggregator sum.

judgments that agent's judgment set differs from the judgment set in the corresponding row. It can be observed from the calculations in the figure, in the furtherst right hand-side column, the set that is closest to all the agents sets is the one contributed by Blue and Orange.

This approach of aggregating the distances between judgment sets, and selecting the set that is at a minimal aggregated distance from the profile, has already been used to design judgment aggregation rules by (Pigozzi, 2006). Distance-based judgment aggregation rules, unlike the rules based on minimization, do not always select the issue-majoritarian judgment set when such exists. However, they are interesting rules to consider for hierarchical groups.

In consensual groups it makes sense to consider that all agents' judgments are of equal weight and importance. If Red, Blue, Green and Orange are members of a consensual group then we need to aggregate their judgments in such a way that each of them has equal bearing on the produced consent. In hierarchical groups, the agent that aggregates the judgments is interested in using the best judgments. What does "best judgments" means? If the agents in the group have different areas and levels of expertise, then it is of advantage to the group to aggregate by considering adequate different weights for different judgments. Lets interpret the agenda issues as follows:

- *a* software upgrades are affordable
- *b* software does not perform according to expectations
- *c* user satisfaction is low
- *d* recommend modification of the software

If agent Red is an expert in finance, then it is better for the group to capitalize on his expertise and assign a high weight for his judgment on *a*. In the presence of weights, the aggregation rule should prioritize higher weights on a judgment and not the number of agents that support it. Furthermore, in the presence of experts, the requirement that all agents express either acceptance or rejection for each agenda issue is rendered meaningless. If an agent is not an expert regarding a particular issue, it is unfeasible to expect and potentially undesirable to request his judgment on this issue. Hence, the agent that aggregates the judgments needs to consider judgment weights, but also rules that can handle incomplete individual judgment sets.

Weighted rules that aggregate incomplete judgment sets have not been proposed in judgment aggregation theory. However, we can extend the distance-based judgment aggregation rules to obtain such rules. Distance-based aggregation rules originate from the theory of belief merging; see for instance the works of (Revesz, 1995; Konieczny and Pino-Pérez, 1999; Konieczny et al., 2004; Condotta et al., 2008). Within belief merging, belief bases are aggregated. In this area of research weights associated with agents are considered, as well as multiple values for the truth-value of the beliefs. We build further on this work in belief merging to construct weighted distance-based rules.

1.2.2 Pairing aggregation contexts with adequate rules

Outside of the domain of law, see (Nash, 2003)[1], very little is known about contexts in which judgment aggregation is, or can be, applied. The contexts of consensual and hierarchical groups are very large and can be further categorized. Each sub-context produces its own desirable properties that a judgment rule applied in it should satisfy.

Within voting contexts, a voting rule is selected based on the properties it satisfies. For instance, the plurality rule, in which each agent chooses one candidate and the candidate with the most votes wins, is used when the number of agents is much larger then the number of candidates and when the agents cannot be expected to spend a lot of time constructing a more complex vote such a total preference order.

Many properties have been proposed for voting rules. These properties are both of structural and relational nature. Structural properties describe the adequate rule based on the structure of the votes, or the desired structure of the winner. Questions considered are such as, do we need strictly one winner or is it acceptable that more than one candidate is selected as winner. The relational properties describe the desirable relations between the profile of individual votes and the winner. For instance, selecting as winner the candidate that is preferred by a majority of agents when compared with any other candidate is a relational property called the *Condorcet winner property*.

In judgment aggregation theory some structural and relational properties are proposed and studied, but not nearly as many as in voting theory. In order to be able to pair aggregation contexts with aggregation rules we need to study the rules we design from the aspect of these properties. We also need to enlarge the set of judgment aggregation rule properties.

The rules we propose are of different structure than the ones considered in the judgment aggregation literature when defining relational and structural properties. Exceptions are the distance-based rules of (Pigozzi, 2006), for which properties have not been extensively studied. Figure 1.3 illustrates the difference between the structures of our rules and those in the literature. The rules for which the properties are defined are partial functions that associate to each profile, of complete judgment sets, one complete judgment set. The rules we define are functions that associate to each profile, a selection of possibly incomplete judgment sets.

Since the output from the judgment aggregation rule is not necessarily a unique judgment set, the relational rules, in general, cannot be directly applied to our rules. Therefore we need to construct new corresponding definitions of these properties.

To enlarge the set of interesting judgment aggregation rule properties, primarily of the relational kind, we consider known properties in voting theory for inspiration. We also define properties only of interest in judgment aggregation, relational properties that consider the relations between the agenda issues, the profile and the selected collective judgment set(s).

[1]Different Nash from the game theory one!

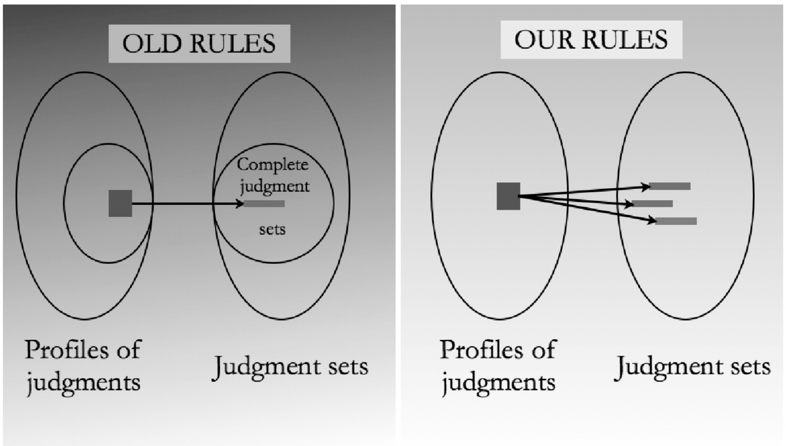


Figure 1.3: Comparing the structure of the rules considered in the literature and those we propose in the thesis.

1.2.3 Applying judgment aggregation rules in MAS

We consider the two general categories of aggregation contexts based on the types of groups that need to make collectively binding decisions. For each of these groups we give an example in which judgment aggregation can be used.

A hierarchical group example

For the example of a hierarchical group consent reaching problem we look for problems in which negotiation cannot be used. We consider a hierarchical team that needs to solve problems in a changing environment.

A decision theoretic approach requires that the group lists all possible options, calculates the expected utilities for each and chooses the option that has the maximal expected utility. When agents are under severe resource restrictions they need to rely on some “approximate” method to make group decisions. However, just as agents do not have the time to negotiate on what the state of the world is, they also do not have the time to enlist and consider all the possible options.

Groups of people, such as teams of firefighters or soldiers on battlefields, do manage to reach consents in uncertain environments despite the resource constraints. We begin by asking if these human decision-making methods can be extended to hierarchical teams of artificial agents. We studied how human teams make decisions, with the purpose of finding a simple model that can be used as a base. Our search yielded the *recognition-primed decision* (RPD) model (Klein et al., 2010).

The RPD model describes how a commander of a firefighting team decides what to do when faced with a familiar problem. This commander does not consider the opinions of the rest of the firefighters when making decisions. According to the RPD model, the commander matches the problem with a typical solution and verifies if the solution is applicable by considering if a particular set of cues are present and/or satisfied. For instance, when faced with a burning house a commander first considers the option of having his team extinguish the fire. This option is adequate if the building is empty of people (first cue) and there are surrounding buildings that are in danger of catching fire themselves (second cue). Since the commander is on site of the burning building he can decide for himself if the cues are present or not. The commander considers the typical solutions one by one until he finds an adequate solution.

We consider a team in which the commander is not on the site where the problem is; there are artificial agents there. For such a team we lift the RPD model to a recognition-primed group decision model (RPgD). The commander determines which is the typical solution and which combination of cues verify its adequacy, but it is the artificial agents that judge if the cues are present or not. The commander uses judgment aggregation to make a final decision whether the solution is adequate and for which reasons, or cues is this decision taken. For the RPgD model we also consider how the group consents on the cues can be used to reconsider group decisions in light of new information.

A consensual group example

As an example of using judgment aggregation for reaching consent in a consensual group, we consider the problem of group intentions. A commonly accepted paradigm is that an agent intends to do something if he chooses it as his goal and is committed to bringing it about (Cohen and Levesque, 1990). But what does it mean for a group to intend something?

Although the problem of what group intentions are, and how they can be represented and generated, has been studied since the nineties. However, the proposed solutions are all based on the same underlying social epistemic theory, that a group intends to do something if and only if all the agents in the group intend individually to the same thing and are committed to doing it. This is the so called *summative view* on what a group attitude is.

The summative group intentions are easy to establish. However, allowing a group to act only when all the agents in the group are of the same mind state poses a limitation on the group's construction and its scope of abilities. It is unfeasible that a group with a lot of members that are heterogeneous, for instance some being robots and others software agents, would be able to align its intentions.

An alternative to the summative view is the so called *non-summative view* of group attitudes. According to this view, a group intends to do something if and only if the members of the group *agree* to do that thing and are jointly committed to doing it. We construct a model of group intentions based on the non-summative view. We propose a procedure for reaching group intentions that relies on a judgment aggregation rule for reaching agreements. The group agrees on what goals to pursue, but also on the reasons for which to pursue that goal. We use these reasons to construct strategies for joint commitment, but also strategies for revising the intentions of the group.

1.3 Interdisciplinarity and methodology

Answering the research question of the thesis calls for an interdisciplinary study among computational social choice, judgment aggregation and multiagent systems. Both computational social choice and judgment aggregation are new disciplines and no textbook or established approach of study exists for either. They, and the field of multiagent systems, are interdisciplinary areas of research in their own right. As a result, the area of study conducted for this thesis spans over social choice theory, social epistemology, experimental psychology and decision making theory in addition to the computer science areas of agent cooperation, agent reasoning, agent modeling and complexity theory.

The sub-problems of designing rules, studying their properties and using them in multiagent systems problems, were not answered sequentially, but rather in parallel, starting with the search for problems of reaching collectively binding decisions in multiagent systems.

Our method for tackling the problem of applying judgment aggregation was to search for both consensual and hierarchical group contexts in which judgment aggregation is a better approach than negotiation.

Negotiation as an agreement reaching technology is difficult to apply in contexts where the agent environment is changing and where the agents are heavily constrained with respect to the time they have to reach consents. Such contexts are usually modeled using a Partially Observable Markov Decision Process (POMDP) (Monahan, 1982). However modeling group decision-making as a POMDP often cannot be solved efficiently. Decision-making is the process of choosing one option from a set of possible options. A rational agent chooses the option that maximizes his expected utility. Simon (1956) prescribed that a resource bound rational agent should not do decision-making at all, since listing the alternative options and calculating the expected utilities is costly. Instead, an agent should *satisfice*, namely he should choose the first option he finds that satisfies a list of sufficient criteria. Satisficing is an appealing approach for groups of agents, however Simon proposed a concept applied to a single agent and not a model for a group.

Human teams of agents such as firefighters, army personal and various emergency rescue teams face uncertain environments and time-constrained decisions. These are also groups in which each agent can be expected to be more reliable on some issues than on others. For instance, a firefighter inside a burning house may be able to observe better if there are victims that need to be rescued than a firefighter that is outside the house. In these teams there is a hierarchical chain of command. The methodology we adopt is to look for decision-making models, constructed by means of experimental psychology, that describe the reasoning of firefighters and various other emergency rescue teams. The cognitive models should be applicable, or modifiable for application to artificial agents.

Judgment aggregation is an adequate method for consensual groups, when they need to reach consents on several, logically related, issues concurrently. An example of a set of related issues is the one consisting of an agent's intention and his beliefs that support and justify his intention. According to (Cohen and Levesque, 1990), an agent's intention is the goal he had chosen to pursue and is committed to pursuing. An agent's choice of a goal is constrained by his beliefs and knowledge of the world. Groups also need to determine their intentions based on what they as a group hold to be true about the world. Beliefs, goals and intentions are called *mental attitudes*.

How collective attitudes are formed and modeled is studied by social epistemology (Goldman, 1987). The above approach to determining the attitudes of a group is often referred to as *summative* (Meijers, 2002). An alternative is modeling the collective attitudes such as intentions in the *non-summative* sense (Gilbert, 2009). According to this approach, a group has a particular intention if the agents *agree* that this is what their intention is. List (2005) proposes that judgment aggregation is used as a formal approach to thinking about institutions in social epistemology. Our methodology is to propose a non-summative model based on judgment aggregation that generates collective intentions from relevant individual beliefs and goals.

We need to develop two categories of rules, one for consensual and one for hierarchical groups. As we observed in the examples of the previous section, adherence to majority, and in general all group decisions that minimize the discrepancy with the individual opinions, are the desirable properties for consensual groups. In voting theory many rules have been proposed based on the concept of minimization: Kemeny, Dodgson, ranked pairs, etc., see (Brams and Fishburn., 2004) for an overview. We also use minimization to develop judgment aggregation rules that are majority adherent.

In hierarchical groups, one agent that aggregates the input from the rest of the group members makes the group decision. This agent does not need to be concerned with having the group decision adherent to the majority. Instead he needs to use the various strengths and competences of the group members, by considering some individual judgments as very relevant and others possibly as useless. The rules suitable for a hierarchical group need to be more general then those for the consensual groups, in the sense that these rules should aggregate the individual judgments by considering weights assigned to the judgments. This category of rules is aimed for the context of uncertain environments. Therefore, in addition to the structural and relational properties, we also need to analyze the complexity-theoretic properties of these rules.

1.4 What is new and what is old

There exists no unique, standard framework of judgment aggregation. In general, one can distinguish between the logic-based frameworks that were introduced by (List and Pettit, 2002) and generalized by (Dietrich, 2007), and abstract or algebraic frameworks introduced by (Wilson, 1975) and extended by (Rubinstein and Fishburn, 1986). Given a logic-based framework, one can construct a corresponding abstract framework. However, for one abstract framework, there are many logic-based frameworks that can be constructed (List and Polak, 2010).

The main differences among frameworks are based on how the agenda issues, the relations among them and the judgments are defined. Some authors, such as (Pauly and van Hees, 2006; Dietrich, 2007; Dietrich and List, 2007b; Dietrich and Mongin, 2010; Endriss et al., 2010a), define the agenda issues as not necessarily atomic formulas of some logic, with the relations among the agenda issues incorporated in the issues themselves. An example of such agenda, using propositional logic and the atoms $\{p,q,r\}$ is $\{p, p \rightarrow q, r \rightarrow p\}$. Other authors, such as (Pigozzi, 2006; Miller, 2008), restrict themselves to atomic agendas and additionally specify a set of constraints that capture the logic relations among the issues. An example of such an agenda is the “doctrinal paradox” originating from (Kornhauser and Sager, 1993; Chapman, 1998). An instance of the “doctrinal paradox” is an agenda: there was a contract (p), assuming there was a contract there is a breach of contract (q), and the defendant is liable for a breach of contract (r). The constraint is $(p \wedge q) \leftrightarrow q$.

Authors like (Dietrich, 2007; Dietrich and Mongin, 2010; Endriss et al., 2010a) require that the agenda is closed under negation, namely that if φ is an issue in the agenda, then so is $\neg\varphi$. These frameworks define a judgment to be a non-empty subset of the $\{\varphi, \neg\varphi\}$ set. Authors such as (Pauly and van Hees, 2006; Pigozzi, 2006; Miller and Osherson, 2009) define a judgment to be a truth valuation of an agenda element. Pauly and van Hees (2006) consider multi-valued truth assignments while (Pigozzi, 2006; Miller and Osherson, 2009) consider strictly binary values. Frameworks exist that impose further constraints on the agenda set, such as for instance that it is closed under deduction. There possibly are contexts in which one framework version is better than another, but these have not been studied or specified in the literature.

We define a general framework for judgment aggregation in which the agenda can contain non-atomic issues, but also additional constrains over the issues can be specified. In this framework, for binary and three-valued judgments, we can express the judgment sets in a dual fashion: as sets of propositions and as sequences of truth-values. We use the logic-based framework of (Dietrich and List, 2007b; Dietrich and Mongin, 2010; Endriss et al., 2010a) for defining the rules based on minimization, but for the rules aimed for hierarchical groups we need the general framework.

The few judgment aggregation rules have been proposed in the literature, most of which were not compared among each other. Comparing two-judgment aggregation rules means determining the relation between the judgments sets produced by each rule. Depending on the structural properties of the collective judgment sets produced, we can define different relations between rules. By comparing rules we primarily are interested in establishing whether two given rules select the same collective judgments or judgment sets for all profiles. We are also interested in which rule selects more judgments or judgment sets for the same profile.

The first collection of judgment aggregation rules, as well as the relations that hold among them and existing rules, were developed as a joint effort with Jérôme Lang, Gabriella Pigozzi and Leendert van der Torre. Part of this work was published in the joint paper (Lang et al., 2011). The full work is a manuscript in preparation for a journal submission. The complexity-theoretic analysis of the second category of rules was a joint effort with Wojciech Jamroga. This analysis, together with the rules themselves, and an analysis of the relationship between these rules and known judgment aggregation rules, was published as (Slavkovik and Jamroga, 2011). An extended version of this paper, including some of the properties of the rules, was submitted to AAMAS 2012. Other published work related to this category of rules and the properties they satisfy, is the joint paper with Gabriella Pigozzi and Leendert van der Torre (Pigozzi et al., 2009).

This thesis is one of the first efforts to develop and implement social choice rules specifically for use in multiagent systems. The difficulty in threading this direction of computational social choice lies first in the lack of unified formalisms in judgment aggregation. Judgment aggregation properties are defined for a particular construction of judgment aggregation rules that does not allow for many rules to be specified, see Figure 1.3. The reputation of social choice theory is that of the theory of impossibility. As insightful and important as the impossibility results are, they do not render the use of social choice rules neither trivial nor impossible.

Algorithmic approaches to applying social choice rules are rare, even in work of computational social choice. Two examples of models for reaching group consent based on judgment aggregation have been developed here, which show how group consent based on judgment aggregation can be implemented. The first example we present is a consent-reaching model for hierarchical groups. It models reaching consent in uncertain environments. This model was constructed by lifting a known cognitive model from an individual agent to a team level. This work is an extension of a paper that was published jointly with Guido Boella, Gabriella Pigozzi, and Leendert van der Torre (Boella et al., 2011b).

The second example we present is a consent reaching model for consensual groups. It models reaching collective intentions in the non-summative social epistemic sense. This model is developed based on the concepts proposed in (Gilbert, 1987, 2002, 2007, 2009). The existence of models such as this has been foreseen in (Dunin-Keplicz and Verbrugge, 2010, Section 3.9). The model of collective intentions has been published as the joint article with Guido Boella, Gabriella Pigozzi, and Leendert van der Torre, (Boella et al., 2011a).

Additional work published work related to this thesis, but not included since it falls outside of the outlined scope of the thesis, is (Pigozzi et al., 2008a,b; Grossi et al., 2009).

1.5 Thesis layout

The thesis is structured as follows. In Chapter 2 the rules based on minimization are presented, as well as the relationships between these rules and the rules that already exist in the judgment aggregation literature. This chapter includes and extends the Sections one, two, three, four, five and seven of (Lang et al., 2011).

In Chapter 3 we present the family of weighted rules for ternary judgments based on distances as well as specific examples of family members. Most of these rules are based on the semantic belief merging operators presented in (Konieczny et al., 2004), but some new rules are introduced as well. The novelty from the belief merging rules is in the introduction of weights for the judgments. We also give a complexity-theoretic analysis of the rule class as a whole. Portions of this chapter were published in (Slavkovik and Jamroga, 2011).

In Chapter 4 we introduce and study properties for judgment aggregation rules. This chapter includes the Section six of (Lang et al., 2011).

Chapters 5 and 6 present the examples of models for hierarchical and consensual groups correspondingly. Chapter 6 predominantly consists of (Boella et al., 2011a), while Chapter 5 is a considerably extended version of (Boella et al., 2011b).

In Chapter 7 we give an overview of related work.

Chapter 8 contains a summary of the thesis and an overview of results. This chapter also includes related work on implementation of judgment aggregation on robots and a set of directions for future work.

Part I

Developing judgment aggregation rules

2

Developing judgment aggregation rules based on minimization

Abstract. Collectively binding decisions by consensual groups, just as winners of democratic elections, need to adhere to the will of the majority, or at least minimize the discrepancy with the opinions held by a majority of the agents. Many voting rules are based on minimization or maximization principles. Likewise, in the field of logic-based knowledge representation and reasoning, many belief change or inconsistency handling operators make use of minimization. The aim of this chapter is to develop and study rules for judgment aggregation based on minimization. We distinguish four families of rules. The rules of the first family first compute the issue-majoritarian judgment set and then restore consistency to this set, when it is inconsistent, using some minimal profile change principle. The rules of the second family proceed in a similar way but take into account the strength of the majority on each issue. Those of the third family consist in restoring the consistency of the majoritarian judgment by removing or changing some individual judgments in a minimal way. Finally, those of the fourth family are based on some predefined distance between judgment sets, and look for a consistent collective judgment minimizing the overall distance to the individual judgment sets. For each family we propose a few typical rules. While most of these rules are new, a few ones correspond to rules that have been defined in the literature. We study the inclusion relationships between these rules to show that they are distinct rules.

2.1 Introduction

In voting theory and in computational social choice, a large body of work focuses on specific voting rules: how their winner sets compare to each other; their social choice-theoretic properties; the computational and communication complexity of winner determination; the theoretical and experimental study of manipulability and control; the amount of information necessary to determine the outcome; *etc.* The focus on specific rules, or families of judgment aggregation rules has been the topic of few papers. We give an overview of these rules.

- The *premise-based* procedure has been introduced in (Kornhauser and Sager, 1993) under the name “issue-by-issue voting and studied in (Dietrich and Mongin, 2010; Mongin, 2008). For this procedure, the agenda is assumed to be partitioned into two subsets: *premises* and *conclusions*. The premises are logically independent. The individuals vote on the premises and the majority on each premise is used to find the collective outcome for that premise. From these collective outcomes on the premises, the collective conclusions are derived using either the logical relationships among, or some external constraints regarding the agenda issues. On the other hand, in the *conclusion-based procedure*, individuals decide privately on the premises and express publicly only their judgments on the conclusions.

- The more general *sequential* procedures (List, 2004a; Dietrich and List, 2007b; Li, 2010) proceed this way: the elements of the agenda are considered sequentially, following a fixed linear order over the agenda (corresponding for instance to temporal precedence or to priority) and earlier decisions constrain later ones. Collective consistency is guaranteed by definition. Of course, in the general case, the result depends on the choice of the order, *i.e.* it is *path-dependent*. Premise-based procedures are specific instances of sequential procedures.

- *Quota-based* rules (Dietrich and List, 2007b; Dietrich, 2010) are a class of rules where each proposition of the agenda is associated with a quota, and the proposition is accepted only if the proportion of individuals accepting it is above the quota. For example, uniform rules take the same quota for all elements of the agenda. The majority rule is a special case of quota-based rules. In Dietrich and List (2007b) sequential quota rules are also considered.

- *Distance-based* rules (Miller and Osherson, 2009; Pigozzi, 2006) assume a predefined distance between judgment sets and/or profiles and choose as collective outcome the consistent judgment sets which are closest (for some notion of closeness) to the individual judgments.

Even if a few families of judgment aggregation rules have been proposed and studied, still the focus on the research is more on the search for impossibility theorems and axiomatic characterizations of families of rules, which contrasts with voting theory, where voting rules are defined and studied *per se*.

In voting theory, quite a number of rules are based on some minimization (or maximization) process: for instance, *Kemeny*, *Dodgson*, *Slater*, *ranked pairs*, *maximin* etc. (We shall not recall the definition of all these voting rules; the reader can refer, for instance, to (Brams and Fishburn., 2004) for a survey.) Minimization is also a common way of defining reasoning rules (such as belief revision operators, inconsistency handling procedures, or nonmonotonic inference rules) in the community of logic-based knowledge representation and reasoning: typically, one deals with inconsistency by looking for maximal consistent subsets of an inconsistent knowledge base. Belief revision often amounts to incorporating a piece of information to a knowledge base while minimizing the information loss from the initial knowledge base. Similar minimization processes are at work in reasoning about action, belief update and belief merging.

In contrast, with the exception of distance-based rules, minimization has rarely been considered for judgment aggregation. Our rules maximize the portion of a profile we wish to keep. The way such maximization is defined depends on the specific rule. Thus, the maximization operated by our aggregation rules is equivalent to minimizing the portion of a profile we wish to remove. In other words, we call our rules “based on minimization”, but we could as well say that our rules are based on maximization. Most of the rules we introduce here are new, while a few of them correspond, up to some minor details, to judgment aggregation rules already proposed in the literature.

From the definitions of two rules, it is sometimes difficult to determine if those rules select the same collective judgments for each profile, if the collective judgments selected by one rule are always also selected by the other. Therefore, when proposing a rule, we need to establish the *inclusion relations* that hold between the new rule and each of the rules already proposed. Not only is this analysis necessary to prove that the rule we are defininig is really new, but it helps in selecting rules for a decision-making problem. Our rules are such that they can select several collective judgment sets for one profile. There are some contexts in which more is better, and other contexts, particularly in multiagent systems require that a unique, or as little as possible, judgment sets are selected.

This chapter is structured as follows. In Section 2.1.1 we introduce the necessary definitions. In Section 2.2 we present the four families of judgment aggregation rules, give examples of specific rules in each family and relate these new specific rules to similar rules in voting theory and/or knowledge representation and reasoning. In Section 2.3 we analyze the inclusion relationships between each pair of introduced rules. In Section 2.4 we make our conclusions and some directions for future work.

2.1.1 General definitions

Let $\mathcal{L}_{prop}$ be a propositional language built on a finite set of propositional symbols $\mathcal{L}_0$. *Cn* denotes logical closure, $Cn(S) = \{\alpha \in L \mid S \models \alpha\}$.

Definition 1 (Agendas, judgment sets, profiles)**.**

- *an* agenda *is a finite set* $\overline{\mathcal{A}} = \{\varphi_1, \neg\varphi_1, \ldots, \varphi_m, \neg\varphi_m\}$ *of formulae of* $\mathcal{L}_{prop}$*, consisting of pairs of propositions* $\varphi_i, \neg\varphi_i$*, where* $\neg\neg\varphi_i \equiv \varphi_i$ *and* $\neg\neg\neg\varphi_i \equiv \neg\varphi_i$*.* $\overline{\mathcal{A}}$ *does not contain tautologies nor contradictions. The* pre-agenda $\mathcal{A}$ *associated with* $\overline{\mathcal{A}}$ *is* $\mathcal{A} = \{\varphi_1, \ldots, \varphi_m\}$*. A* subagenda *of* $\overline{\mathcal{A}}$ *is a subset of an agenda* $\overline{\mathcal{A}}$ *that also contains pairs of propositions* $\varphi, \neg\varphi$*, where* $\varphi \in \mathcal{A}$*.*
- *a* judgment set *J over* $\overline{\mathcal{A}}$ *is a subset of* $\overline{\mathcal{A}}$*. A judgment set J is* complete *if for every pair* $\{\varphi, \neg\varphi\} \subseteq \overline{\mathcal{A}}$*, J contains either* φ *or* $\neg\varphi$*. A judgment set J is* consistent *if it is a satisfiable set in terms of classical propositional logic. The set* $\hat{\mathbb{A}}(\overline{\mathcal{A}})$ *is the set of all consistent non-empty judgment sets that can be constructed over* $\overline{\mathcal{A}}$*. The set* $\Phi(\overline{\mathcal{A}})$ *is the set of all* complete *and consistent judgment sets that can be constructed over* $\overline{\mathcal{A}}$*.*
- *an n-*voter profile *over* $\overline{\mathcal{A}}$ *is a collection* $P = \langle J_1, \ldots, J_n \rangle$ *where each* J_i *is a consistent and complete judgment set.*

We now define judgment aggregation rules. We write $\hat{\mathbb{A}}$ shortly for $\hat{\mathbb{A}}(\overline{\mathcal{A}})$ and Φ shortly for $\Phi(\overline{\mathcal{A}})$ to improve readability.

Definition 2 (Judgment aggregation rules)**.**

- *A* deterministic judgment aggregation rule *is a function* $f_{n,\overline{\mathcal{A}}} : \Phi^n \mapsto \Phi$*. Namely,* $f_{n,\overline{\mathcal{A}}}$ *associates with every profile* $P = \langle J_1, \ldots, J_n \rangle$ *a consistent and complete judgment set* $f_{n,\overline{\mathcal{A}}}(P)$*.*

- *an* irresolute judgment aggregation rule *(or* judgment aggregation correspondence*) is a function* $F_{n,\overline{\mathcal{A}}} : \Phi^n \mapsto \mathcal{P}(\hat{\mathbb{A}})$, *associating with every profile P a nonempty set of consistent, possibly incomplete, judgment sets* $F_{n,\overline{\mathcal{A}}}(P)$.

Often in the judgment aggregation literature, a judgment aggregation rule is defined as a function whose co-domain is the full set $\{0,1\}^m$, where m is the agenda cardinality. The requirement that the rule produces only consistent judgment sets is additionally specified as a property, called *consistency*, of the judgment aggregation function. Here we opted for the definition that specifies Φ as a domain, as consistency is not a property that we can compromise on while applying judgment aggregation in multiagent systems. This point will be further clarified in Chapters 5 and 6.

Most of the time, when referring to judgment aggregation rules we will keep n and $\overline{\mathcal{A}}$ implicit when they are clear from the context, *i.e.*, $f_{n,\overline{\mathcal{A}}}$ (resp. $F_{n,\overline{\mathcal{A}}}$) will be simply denoted as f (respectively F). Also, by a slight abuse of language, if $P = \langle J_1,\ldots,J_n\rangle$, then we will write $f(J_1,\ldots,J_n)$ and $F(J_1,\ldots,J_n)$ instead of $f(\langle J_1,\ldots,J_n\rangle)$ and $F(\langle J_1,\ldots,J_n\rangle)$.

As in voting theory, a rule can be obtained from a correspondence using a tie-breaking mechanism, such as a priority over judgment sets, or over agents. In this chapter we focus on irresolute rules, unless we state the contrary.

There are two different views of aggregation rules: either we see the output as a mere collection of consistent judgment sets, or we see it as a closed logical theory.

Definition 3 (Logical theory $T_F(P)$)**.** *Given a judgment aggregation rule F, and a profile P, we define the logical theory* $T_F(P) = \bigcap\{Cn(J) \mid J \in F(P)\}$.

Definition 4 (Rule equivalence)**.** *Let F and F′ be two aggregation rules. F and F′ are* theory equivalent, *denoted* $F =_T F'$ *if for every profile P we have* $T_F(P) = T_{F'}(P)$. *F and F′ are* equal, *denoted* $F = F'$, *if for every profile P we have that* $F(P) = F(P')$.

Definition 5 (Rule inclusion)**.** *Let F and F′ be two aggregation rules. F is* at least as discriminant *as F′ if for every profile P we have* $T_{F'}(P) \subseteq T_F(P)$. *F and F′ are* incomparable *if there exist two profiles P and Q such that* $T_F(P) \nsubseteq T_{F'}(P)$ *and* $T_F(Q) \nsubseteq T_{F'}(Q)$.

A formula α is in $T_F(P)$ if and only if it can be inferred from every judgment set in $F(P)$. Note that $T_F(P)$ being the intersection of consistent closed logical theories, is itself a consistent closed theory. Intuitively, F is at least as discriminant as F' if, for every profile P, all judgments included in every set $F'(P)$ are necessarily included in every set $F(P)$.

Definition 6 (Majoritarian aggregation)**.** *Let* $\varphi \in \overline{\mathcal{A}}$. *The issue-majority aggregation rule m is a function* $m : \Phi^n \times \{\varphi,\neg\varphi\}^n \mapsto \{\varphi,\neg\varphi,\varnothing\}$ *defined as:*

$$m(P,\varphi) = \begin{cases} \varphi & \textit{iff}\ \#\{i|\varphi \in J_i\} > \frac{n}{2} \\ \neg\varphi & \textit{iff}\ \#\{i|\neg\varphi \in J_i\} > \frac{n}{2} \\ \varnothing & \textit{iff}\ \#\{i|\varphi \in J_i\} = \#\{i|\neg\varphi \in J_i\} \end{cases}$$

$M(P)$ is a judgment set defined as $M(P) = \{m(P,\varphi) \mid \varphi \in \overline{\mathcal{A}}\}$.

$M(P)$ is the *issue-majoritarian judgment set* associated with P. Note that $M(P)$ is not necessarily an element of Φ, nor of $\hat{\mathbb{A}}$.

In the remainder of this chapter we call the set $M(P)$ the majoritarian set.

Definition 7 (Majority-preservation)**.** *A profile P is* majority-consistent *if $M(P) \in \hat{\mathbb{A}}$. A judgment aggregation rule F is* majority-preserving *if, for every majority-consistent profile P, $F(P) = M(P)$.*

If n is odd then $M(P)$ is necessarily a complete judgment set.

Example 2.1.1. *Consider the pre-agenda $\mathcal{A} = \{p \wedge r, p \wedge s, q, p \wedge q, t\}$ and a profile P of 17 voters, presented in Table 2.1.*

Voters	$\{\ p \wedge r,$	$p \wedge s,$	$q,$	$p \wedge q,$	$t\ \}$
$J_1 \times 6$	+	+	+	+	+
$J_2 \times 4$	+	+	-	-	+
$J_3 \times 7$	-	-	+	-	-
M(P)	+	+	+	-	+

Table 2.1: The profile P

We obtain $M(P) = \{p \wedge r, p \wedge s, q, \neg(p \wedge q), t\}$. $M(P)$ is an inconsistent judgment set, therefore P is not majority-consistent.

We end this Section by defining distances between judgment sets and between a judgment set and a profile. A distance d between judgment sets over $\overline{\mathcal{A}}$ is a function $d : \Phi \times \Phi \rightarrow \mathbb{N}^0$ such that for all $J, J', J'' \in \Phi$:

(a) $d(J, J') = 0$ if and only if $J = J'$,

(b) $d(J, J') = d(J', J)$, and

(c) $d(J, J'') \leqslant d(J, J') + d(J', J'')$.

A distance function between profiles is defined similarly. Finally, the *Hamming distance* between judgment sets and between profiles (Miller and Osherson, 2009; Endriss et al., 2010b) is defined as follows.

Definition 8 (Hamming distance between complete judgment sets d_H)**.** *Given two complete judgment sets J and J' (over the same set of agents and the same agenda), the Hamming distance d_H between J and J' is defined by*

$$d_H(J, J') = |J \backslash J'| + |J' \backslash J|$$

Now, the distance between two profiles is the sum of the Hamming distances between their individual judgment sets.

Definition 9 (Hamming distance between profiles D_H)**.** *Given two profiles $P = \langle J_1, \ldots, J_n \rangle$ and $Q = \langle J'_1, \ldots, J'_n \rangle$, the Hamming distance between P and Q is defined by*

$$D_H(P, Q) = \sum_{i=1}^{n} d_H(J_i, J'_i)$$

Instead of the sum, another algebraic aggregation function D can be used as well, such as for instance *min*, *max* or Π, under the conditions $(a) - (c)$ are observed by D.

2.2 Four families of aggregation rules

In this section we present four different families of minimization-based judgment aggregation rules.

2.2.1 Rules based on the majoritarian judgment set

Definition 10 (Rule based on the majoritarian judgment set)**.**

A rule R is based on the majoritarian judgment set *if there exists a function g mapping every judgment set (consistent or not) to a nonempty set of consistent judgment sets, such that for every profile P,* $R(P) = g(M(P))$*.*

This family can be viewed as the judgment aggregation counterpart of voting rules that are based on the pairwise majority graph, also known as tournament solutions. Being based on the majoritarian judgment set means that knowing the majoritarian judgment set of a profile is enough to determine the outcome of the rule. Equivalently, two profiles P and Q whose majoritarian judgment coincide ($M(P) = M(Q)$) will lead to the same outcome ($R(P) = R(Q)$). We naturally expect these rules to be majority-preserving, which is equivalent to saying that the restriction of g to consistent judgment sets is the identity: if J is consistent, then $g(J) = \{J\}$; such a condition can be seen as the counterpart, for judgment aggregation, of Condorcet-consistency.

When $M(P)$ is not consistent, we look for a minimal way of restoring consistency by removing some elements from the agenda. Given a judgment set J, we define the set of *consistent sub judgment sets of J*, denoted by $Cons(J)$, as $g(J) = \{J' \subseteq J \mid J' \in \Phi\}$. Defining a rule consists in defining a minimalism criteria for the set of formulas removed from J. There are two obvious choices, consisting in choosing consistent sub judgment sets of $M(P)$ that are maximal for, respectively, set inclusion or cardinality, which corresponds to the following choices for g:

- $g(J) = \max(Cons(J), \subseteq)$;
- $g(J) = \max(Cons(J), |.|)$

Equivalently, these rules consist in looking for a minimal subset of formulas in $\overline{\mathcal{A}}$ to remove such that the profile becomes majority-consistent. We give a formal definition that corresponds to this alternative characterization.

In the following we use the abbreviation *maxcard* for *of maximal cardinality*.

Definition 11 (Maximal sub-agenda rule R_{MSA})**.** *Given a profile* $P = \langle J_1, \ldots, J_n \rangle$ *on an agenda* $\overline{\mathcal{A}}$*,* $\mathcal{A}$ *the preagenda associated with* $\overline{\mathcal{A}}$*, and a sub-preagenda* $[Y] \subseteq \mathcal{A}$*, the restriction of P to Y is* $P^{\downarrow Y} = \langle J_j \cap Y, 1 \leqslant j \leqslant n \rangle$*. Let* $MSA(P)$ *be the set of all maximal sub preagendas* $[Y]$ *of* $\mathcal{A}$ *(with respect to set inclusion) such that* $P^{\downarrow Y}$ *is majority-consistent. The* maximal sub-agenda *judgment aggregation rule* R_{MSA} *maps P to* $R_{MSA}(P) = \{M(P^{\downarrow Y}) \mid [Y] \in MSA(P)\}$*.*

Example 2.2.1. *Consider the same agenda and profile as in Example 2.1.1. We obtain that*

$$R_{MSA}(P) = \left\{ \begin{array}{llll} \{p \wedge r, & p \wedge s, & q, & t\}, \\ \{p \wedge r, & p \wedge s, & \neg(p \wedge q), & t\}, \\ \{q, & \neg(p \wedge q), & & t\} \end{array} \right\}.$$

Instead of looking for maximal majority-consistent sub-agendas with respect to inclusion we may instead look for maxcard majority-consistent sub-agendas, which leads to the following judgment aggregation rule.

Definition 12 (Maxcard sub-agenda rule R_{MCSA})**.** *Let $MCSA(P)$ the set of all maxcard sub-preagendas $[Y]$ of $\mathcal{A}$ such that $P^{\downarrow Y}$ is majority-consistent. The* maxcard sub-agenda *judgment aggregation rule R_{MCSA} maps P to $R_{MCSA}(P) = \{M(P^{\downarrow Y}) \mid [Y] \in MCSA(P)\}$.*

Example 2.2.2. *Consider again the agenda and profile from Example 2.1.1. The sub-preagenda Y which gives a majority-consistent $P^{\downarrow Y}$ and is maximal is obtained for either $Y = \{p \wedge r, p \wedge s, q, t\}$ or $Y = \{p \wedge r, p \wedge s, p \wedge q, t\}$. We obtain*

$$R_{MCSA}(P) = \left\{ \begin{array}{l} \{p \wedge r,\ p \wedge s,\ \qquad\quad q,\ t\}, \\ \{p \wedge r,\ p \wedge s,\ \neg(p \wedge q),\ \ t\} \end{array} \right\}.$$

The R_{MCSA} rule corresponds, up to some minor details and for a specific choice of a distance function, namely the Hamming distance d_H, to the ENDPOINT judgment aggregation rule defined in (Miller and Osherson, 2009). According to the ENDPOINT rule, the collective judgment sets for P are the consistent and complete judgment sets that are at a minimal distance d from $M(P)$.

We show that the rules R_{MSA} and R_{MCSA} are based on the majoritarian judgment set.

Proposition 2.2.3. *Let $Cons(M(P))$ be the set of all consistent subsets of $M(P)$.*

- $R_{MSA}(P) = \max(Cons(M(P)), \subseteq)$.
- $R_{MCSA}(P) = \max(Cons(M(P)), |.|)$.

Proof. We give the only the proof for R_{MSA}. The proof for $R_{MCSA}(P)$ proceeds exactly in the same way.

Let $[Y] \in MSA(P)$. We have $M(P^{\downarrow Y}) \subseteq M(P)$ and $M(P^{\downarrow Y})$ is consistent. Assume that $M(P^{\downarrow Y})$ is not a maximal consistent subset of $M(P)$. There exists a consistent sub-agenda Z of $\overline{\mathcal{A}}$ such that $M(P^{\downarrow Y}) \subset Z \subseteq M(P)$. Since both $M(P^{\downarrow Y})$ and Z contains at most one of $\varphi, \neg\varphi$ for every $\varphi \in \overline{\mathcal{A}}$ (otherwise they would not be consistent), there must be a φ such that either $\varphi \in Z$ or $\neg\varphi \in Z$, and $\varphi \notin [Y]$. But then $[Y] \cup \{\varphi\} \subseteq Z \subseteq M(P)$ and Z consistent implies that $M([Y] \cup \{\varphi\})$ is a consistent subset of $M(P)$, contradicting $[Y] \in MSA(P)$. Therefore, $M(P^{\downarrow Y}) \in \max(Cons(M(P)), \subseteq)$.

Conversely, let $Z \in \max(Cons(M(P)), \subseteq)$. $Y = \{\varphi \in \overline{\mathcal{A}} \mid \varphi \in Z \text{ or } \neg\varphi \in Z\}$ is a preagenda of $\overline{\mathcal{A}}$, and because Z is a consistent subset of $M(P)$, Z contains at most one of $\varphi, \neg\varphi$ for every $\varphi \in \overline{\mathcal{A}}$, therefore $M(P)^{\downarrow Y} = Z$.

Assume there is a $Y' \supset Y$ such that $M(P^{\downarrow Y'})$ is consistent. Then $M(P^{Y'}) \supset M(P^{Y}) = Z$, contradicting $Z \in \max(Cons(M(P)), \subseteq)$. Therefore, Y is a maximal consistent sub preagenda of P. □

We note that even when n is odd, $R_{MSA}(P)$ and $R_{MCSA}(P)$ may contain incomplete judgment sets. Take for instance $P = \langle \{a, b, a \wedge b\}, \{a, \neg b, \neg(a \wedge b)\}, \{\neg a, b, \neg(a \wedge b)\} \rangle$; then $R_{MSA}(P) = R_{MCSA}(P) = \{\{a, b\}, \{a, \neg(a \wedge b)\}, \{b, \neg(a \wedge b)\}\}$. However, when n is odd, every judgment set in $R_{MSA}(P)$ and *a fortiori* in $R_{MCSA}(P)$ is *equivalent* to a complete judgment set: here, $\{a, b\}$, $\{a, \neg(a \wedge b)\}$ and $\{b, \neg(a \wedge b)\}$ are equivalent to, respectively, $\{a, b, a \wedge b\}$, $\{a, \neg b, \neg(a \wedge b)\}$ and $\{\neg a, b, \neg(a \wedge b)\}$.

Proposition 2.2.4. *If n is odd then for every $J \in R_{MSA}(P)$ and every $J \in R_{MCSA}(P)$, there is a complete judgment set J' such that the deductive closures of J and J' are equivalent.*

Proof. Let $J \in R_{MSA}(P)$ and assume J is not equivalent to a complete judgment set. There is a $\varphi \in \overline{\mathcal{A}}$ such that neither $J \models \varphi$ nor $J \models \neg\varphi$. Because n is odd, $M(P)$ is a complete judgment set, and contains either φ or $\neg\varphi$. Without loss of generality, assume it contains φ. Then $J \cup \{\varphi\} \subseteq M(P)$ and $J \cup \{\varphi\}$ is consistent, contradicting $J \in R_{MSA}(P)$. The proof for R_{MCSA} follows from the fact that $R_{MCSA}(P) \subseteq R_{MSA}(P)$. □

While, as far as we are aware, R_{MSA} is new, R_{MCSA} coincides with the $\texttt{Endpoint}_d$ rule defined in (Miller and Osherson, 2009). We repeat the definition here using our terminology. Recall that Φ is the set of all complete and consistent judgment sets for $\overline{\mathcal{A}}$.

Definition 13 (Endpoint rule)**.** *Let d be a distance function between judgment sets. The judgment aggregation rule $\texttt{Endpoint}_d$ is defined as:*

$$\texttt{Endpoint}_d(P) = \{J \in \Phi \mid d(J, M(P)) \leqslant d(J', M(P)) \text{ for all } J' \in \Phi\}$$

Proposition 2.2.5. $R_{MCSA} =_T \texttt{Endpoint}_{d_H}$.

Proof. A judgment set $J \in \Phi$ extends a judgment set $A \in \hat{\mathbb{A}}$, alternatively J is an extension of A, when if $\varphi \in A$, then $\varphi \in J$. We claim that for every $P \in \Phi^n$ and every $A \in \hat{\mathbb{A}}$, we have $A \in MSA(P)$ if and only if for every $J \in \Phi$ extending A, and every $J' \in \Phi$ it holds that $d_H(J, M(P)) \leqslant d_H(J', M(P))$.

We show the fist direction. Assume that A is a consistent subset of $M(P)$ and let $J \in \Phi$ be an extension of A. We have that $d_H(J, M(P)) \leqslant m - |A|$. We need to show that for every $J' \in \Phi$ it holds that $d_H(J, M(P)) \leqslant d_H(J', M(P))$.

Assume the contrary, namely that there exists $J' \in \Phi$ such that $d_H(J', M(P)) < d_H(J, M(P)) \leqslant m - |A|$. $|J' \cap M(P)| > |A|$ and $J' \cap M(P)$ is a consistent subset of $M(P)$. As a consequence $A \notin MSA(P)$. We conclude that for every $J \in \Phi$ extending A it is the case that $d_H(J, M(P)) \leqslant d_H(J', M(P))$ for every $J' \in \Phi$. Therefore, $T_{\texttt{Endpoint}}(P) \models T_{R_{MCSA}}(P)$.

We show the second direction. Assume that $J \in \Phi$ and $A = J \cap M(P)$. We have that $d_H(J, M(P)) = m - |A|$ and A is a consistent subset of $M(P)$. We need to show that $A \in MSA(P)$.

Assume the contrary, namely that $A \notin MSA(P)$. If $A \notin MSA(P)$, then there exists a consistent subset A' of $M(P)$ such that $|A'| > |A|$. But now, any $J' \in \Phi$ extending A' is such that $d_H(J', M(P)) \leqslant m - |A'| < m - |A| = d_H(J, M(P))$, which implies that we do not have $d_H(J, M(P)) \leqslant d_H(J', M(P))$ for every $J' \in \Phi$. Therefore, $d_H(J, M(P)) \leqslant d_H(J', M(P))$ for every $J' \in \Phi$ implies that $J \cap M(P) \in MSA(P)$. We conclude that $T_{R_{MCSA}}(P) \models T_{\texttt{Endpoint}}(P)$. □

2.2.2 Rules based on the weighted majoritarian judgment set

We first define the *weighted majoritarian judgment set* of a profile P as

$$w(P) = \{\langle \varphi, N(P, \varphi)\rangle, \varphi \in \overline{\mathcal{A}}\}$$

where $N(P, \varphi) = \#\{i, \varphi \in J_i\}$.

Whereas $M(P)$ keeps only the information about which one, of the two propositions φ and $\neg\varphi$, is supported by a majority of voters, $w(P)$ keeps much more information, since it stores the number of voters who support φ and $\neg\varphi$. The set $M(P)$ can be recovered from $w(P)$, but not *vice versa*.

Definition 14 (Rule based on the weighted majoritarian judgment set)**.**

A rule R is based on the weighted majoritarian judgment set *if there exists a function g mapping every judgment set (consistent or not) to a nonempty set of consistent judgment sets, such that for every profile P,* $R(P) = g(w(P))$.

This family can be viewed as the judgment aggregation counterpart of voting rules that are based on the weighted pairwise majority graph, such as maximin, ranked pairs, or Borda.

The first rule of this class we consider is the *maxweight sub-agenda rule*.

R_{MSA} and R_{MCSA} consider the judgments on the agenda subset as a unit that is to be kept in its entirety or got ridden of. A finer way of defining a judgment rule consists in looking for maximal or maxcard majority-consistent subsets of the set of *elementary pieces of information* consisting each of a pair (element of the agenda, judgment on it elicited from an agent). Equivalently, this comes down to weigh each element of the agenda by the number of agents supporting it, and then to look for maxweight sub-agendas.

Definition 15 (Maxweight sub-agenda rule R_{MWA})**.** *For any sub-agenda* $Y \subseteq \overline{\mathcal{A}}$*, the weight of Y with respect to P is defined by* $w_P(Y) = \sum_{\psi \in Y} N(P,\psi)$*. Let MWA(P) be the set of all consistent sub-agendas Y of* $\overline{\mathcal{A}}$ *maximizing* w_P*. The* maxweight sub-agenda *judgment aggregation rule* R_{MWA} *maps P to* $R_{MWA}(P) = \{Y \mid Y \in MWA(P)\}$.

Example 2.2.6. *Consider the agenda and profile of Example 2.1.1. We obtain:*

$N(P, p \wedge r) = 10,\ \ N(P, \neg(p \wedge r)) = 7$
$N(P, p \wedge s) = 10,\ \ N(P, \neg(p \wedge s)) = 7$
$N(P,q) = 13,\ \ N(P, \neg q) = 4$
$N(P, p \wedge q) = 6,\ \ N(P, \neg(p \wedge q)) = 11$
$N(P,t) = 10,\ \ N(P, \neg t) = 7$

$R_{MWA}(P) = \{\{p \wedge r, p \wedge s, q, p \wedge q, t\}\}$*, due to* $w_P(\{p \wedge r, p \wedge s, q, p \wedge q, t\}) = 49$ *is maximal with respect to all other complete and consistent* $Y \subset \overline{\mathcal{A}}$.

The intuition behind this rule is that we look for a minimal number of *elementary information items* to remove from P so that it becomes majority-consistent. An information item is an element from $\overline{\mathcal{A}}$ approved by an agent. The set of information items associated with P, denoted by $\Sigma(P)$, is the multiset containing as many occurrences of φ as agents approving φ in P. *E.g.*, if $\mathcal{A} = \{a,b,c,a \wedge b\}$ and $P = \langle\{a,b,c,a \wedge b\}, \{\neg a,b,c,\neg(a \wedge b)\}, \{a,\neg b,c,\neg(a \wedge b)\}\rangle$, then $\Sigma(P) = \{a,\ b,\ c,\ a \wedge b,\ \neg a,\ b,\ c,\ \neg(a \wedge b),\ a,\ \neg b,\ c,\ \neg(a \wedge b)\}$ and $\Sigma(P) = \{a,a,\neg a,b,b,\neg b,c,c,c,a \wedge b,\neg(a \wedge b),\neg(a \wedge b)\}$.

Let $MaxCard(\Sigma(P))$ be the set of all maxcard consistent subsets of $\Sigma(P)$. If $S \in MaxCard(\Sigma(P))$, then for every $\varphi \in \overline{\mathcal{A}}$, S contains either all occurrences of φ in $\Sigma(P)$ or all occurrences of $\neg\varphi$ in $\Sigma(P)$. Let J_S be the judgment set containing φ if S contains all occurrences of φ in $\Sigma(P)$ and $\neg\varphi$ if S contains all occurrences of $\neg\varphi$ in $\Sigma(P)$. Then $R_{MWA}(P) = \{J_S \mid MaxCard(\Sigma(P))\}$.

Although it looks entirely new, we will show that this natural rule corresponds to a rule already defined, in a different way, in (Endriss et al., 2010b).

Proposition 2.2.7. *R_{MWA} is majority-preserving.*

Proof. Let P be a majority-consistent profile. We claim that $R_{MWA}(P)$ consists of all complete consistent sub-agendas extending $M(P)$. When n is odd, then $M(P)$ is a complete sub-agenda, so in this case $R_{MWA}(P) = M(P)$. However, if n is even then $M(P)$ might be incomplete. For instance, if $n = 2$, $\mathcal{A} = \{p,q\}$ and $P = \langle\{p,q\},\{p,\neg q\}\rangle$ then $M(P) = \{p\}$ and $R_{MWA}(P) = \{\{p,q\},\{p,\neg q\}\}$. Let J be a complete consistent judgment set extending $M(P)$. If $J \notin R_{MWA}(P)$, then there exists a consistent judgment set J' such that $\sum_{\varphi\in J'} N(P,\varphi) > \sum_{\varphi\in J} N(P,\varphi)$. This implies that there must be a $\varphi \in \overline{\mathcal{A}}$ such that $\varphi \in J$, $\neg\varphi \in J'$, and $N(P,\neg\varphi) > N(P,\varphi)$. The latter implies that $\varphi \notin M(P)$, which contradicts the assumption that J extends $M(P)$. □

The following rule is inspired from the ranked pairs rules in voting theory (Schulze, 2003). It consists in first fixing the truth value for the elements of the agenda with the largest majority. It proceeds to iterate, considering the elements of the agenda in the decreasing order of the number of agents who support them, and fix each agenda issue value to the majoritarian value. It proceeds iterating as long as this is possible without producing an inconsistency.

Definition 16 (Ranked agenda R_{RA}). *Let $Y = \{\varphi \in \overline{\mathcal{A}} | N(P,\varphi) > \frac{n}{2}\}$, and let $\geq_P$ be the complete weak order relation on Y defined by $\varphi \geq_P \psi$ if $N(P,\varphi) \geqslant N(P,\psi)$. $R_{RA}(P)$ is defined as follows: $J \in R_{RA}(P)$ if there exists a linear order $\succ_P$ on $\overline{\mathcal{A}}$ refining $\geqslant$ such that $RA(\succ,P) = J$, where $RA(\succ,P)$ is defined inductively by*

- *order the elements of Y using the relation $\succ$, i.e., such that $\varphi_{\sigma(1)} \succ \ldots \succ \varphi_{\sigma(m)}$;*
- *$D := \emptyset$;*
- *for $k := 1$ to m do: if $D \cup \{\varphi_{\sigma(i)}\}$ is consistent then $D := D \cup \{\varphi_{\sigma(i)}\}$;*
- *$RA(\succ,P) := D$.*

R_{RA} is based on the weighted majoritarian judgment set. This rule is a special case of the sequential aggregation rules defined by (List, 2004b), where the rule is defined without specifying a particular order of aggregation.

Example 2.2.8. *Consider the same profile as in Example 2.1.1. We have $Y = \{p \wedge r, p \wedge s, q, \neg(p \wedge q), t\}$, and $q \succ_P \neg(p \wedge q) \succ_P p \wedge r \sim_P p \wedge s \sim_P t$ (where $\sim_P$ and $\succ_P$ are respectively the indifference and the strict preference relations induced from $>$). We obtain*

$$R_{RA}(P) = \{\{q, \neg(p \wedge q), t, \neg(p \wedge r), \neg(p \wedge s)\}\}.$$

Every judgment set J in $R_{RA}(P)$ is complete; if not, there would be a $\varphi \in \overline{\mathcal{A}}$ such that neither φ not $\neg\varphi$ is in J. Since J is consistent, either $J \cup \{\varphi\}$ or $J \cup \{\neg\varphi\}$ is consistent. But then, either φ or $\neg\varphi$ would have been incorporated in J, which contradicts the assumption that J contains neither φ nor $\neg\varphi$. More generally, when the number of voters n is odd, each of the collective judgment sets obtained from any of the rules introduced so far is equivalent to a complete judgment set.

Proposition 2.2.9. *R_{RA} is majority-preserving.*

Proof. In $>$, the elements of $M(P)$ are considered before the elements of $\overline{\mathcal{A}}\backslash M(P)$. Therefore, when an element φ of $M(P)$ is considered, the current judgment set D is a subset of $M(P)$ and $D \cup \{\varphi\} \subseteq M(P)$, therefore $D \cup \{\varphi\}$ is consistent, which implies that φ is incorporated into D. Since this is true for any $\varphi \in \overline{\mathcal{A}}$, we get that any element of $R_{RA}(P)$ contains $M(P)$.

Now, let J be a consistent, complete extension of $M(P)$. Take $>$ such that all elements of $M(P)$ are considered first, then all elements of $J\backslash\{M(P\}$, then all elements of $\overline{\mathcal{A}}\backslash J$. This order refines $>_P$, because if $\varphi \in PM(P)$ then $N(P,\varphi) > \frac{n}{2}$, if $\varphi \in J\backslash\{M(P\}$ then $N(P,\varphi) = \frac{n}{2}$ and if $\varphi \in \overline{\mathcal{A}}\backslash J$ then $N(P,\varphi) \leqslant \frac{n}{2}$. Lastly, $RA(>,P) = J$, which proves that $J \in R_{RA}(P)$. □

2.2.3 Rules based on the removal or change of individual judgments

The principle at work, for this family, is that we look for a modified profile, as close as possible to the original profile (with respect to a given distance), such that the resulting profile is majority-consistent. Different rules will be obtained with different distance functions.

This family can be viewed as the judgment aggregation counterpart of voting rules that are based on performing minimal operations on profiles with the purpose of obtaining a profile for which a Condorcet winner exists. Such are the Young (Young, 1995) and Dodgson rules (Dodgson, 1876). See (Elkind et al., 2009) for a general family of voting rules of that kind.

The first rule we consider is called the *Young* rule for judgment aggregation, by analogy with the Young rule in voting, which outputs the candidate x minimizing the number of voters to remove from the profile so that x becomes a Condorcet winner.

Definition 17 (Young rule for judgment aggregation R_Y)**.**

Given a profile $P = \langle J_1,\dots,J_n\rangle$ and a subset of agents $N^* \subseteq \{1,\dots,n\}$, the restriction of P to N^* is $P^- = \langle J_j, j \in N^*\rangle$, and is called a sub profile *of P. Let $MSP(P)$ be the set of maxcard majority-consistent sub profiles of P for which $M(P^-)$ is a complete judgment set. Then the* Young *judgment aggregation rule Y maps P to $R_Y(P) = \{M(P^-) \mid P^- \in MSP(P)\}$.*

Intuitively, this rule consists of removing a minimal number of agents so that the profile becomes majority-consistent. Or, equivalently, we maximize the number of voters we keep of a given profile. If the profile P is majority-consistent, then no voter needs to be removed and $Y(P) = \{M(P)\}$, hence Y is majority-preserving.

Example 2.2.10. *Once again we consider P for $\overline{\mathcal{A}}$ given in Example 2.1.1. The result*

$$R_Y(P) = \{\{\neg(p \wedge r), \neg(p \wedge s), q, \neg(p \wedge q)\}\}$$

is obtained by removing 3 of the judgment sets $\{p \wedge r, p \wedge s, q, (p \wedge q), t\}$. Removing less judgment sets, or other 3 judgment sets, does not lead to a majority-consistent profile.

Now, instead of looking for a minimal *number* of individual judgments to remove, we can look for a minimal *set* of individual judgments to remove, leading to the following rule.

Definition 18 (Inclusion-Young rule R_{IY})**.**

Let $mSP(P)$ be the set of maximal majority-consistent sub profiles P^- of P. Then the Inclusion-Young *judgment aggregation rule Y maps P to*

$$R_{IY}(P) = \{M(P^-) \mid P^- \in mSP(P)\}$$

Example 2.2.11. *The inclusion Young rule applied to the the profile P results in*

$$R_{IY}(P) = \left\{ \begin{array}{lllll} \{\neg(p \wedge r), & \neg(p \wedge s), & q, & \neg(p \wedge q) & \}, \\ \{p \wedge r, & p \wedge s, & q, & p \wedge q, & t\} \end{array} \right\}.$$

R_{IY} is majority-preserving for the same reason as R_Y.

If a profile is not majority-preserving then one might look at the problem from a different view point. The Young rule and the Inclusion Young rule minimally alter the profile by removing agent's judgment sets. The profile can be minimally altered also by repeating agent's judgment set, extending the profile instead of shrinking it. The intuition behind extending is the assumption that there is confirmation pending for some of the judgment sets in the profile. We may ask which is the least amount of confirmation, *i.e.*, what is the smallest super-profile of the majority-inconsistent profile P that is itself majority-consistent. We thus obtain a new rule, the *reversed Young judgment aggregation rule*, which is also majority-preserving.

Definition 19 (Reversed Young rule for judgment aggregation R_{RY})**.** *Let $P = \langle J_1, \ldots, J_n \rangle$ be a profile. A* super profile *of P is a profile $P^+ = \langle J_1, \ldots, J_q \rangle$, where $q \geqslant n$, such that for every $i \in n+1, \ldots, q$ there exists a $j \leqslant n$ such that $J_i = J_j$. Let $MSA(P)$ be the set of minimal (with respect to cardinality) majority-consistent super profiles P^+ of P. The reverse Young judgment aggregation rule RY maps P to $R_{RY}(P) = \cup\{R_{MSA}(P^+) \mid P_S \in MSA(P^+)\}$.*

Example 2.2.12. *For the profile P of Example 2.1.1, the outcome*

$$R_{RY}(P) = \{\{\neg(p \wedge r), \neg(p \wedge s), q, \neg(p \wedge q)\}\}$$

is obtained by adding 3 of the judgment sets $\{\neg(p \wedge r), \neg(p \wedge s), q, \neg(p \wedge q), \neg t\}$. Adding less, or other 3 judgment sets, does not lead to a majority-consistent profile.

Comparing Examples 2.2.12 and 2.2.10, we observe that $R_Y(P) = R_{RY}(P)$. However, this is not the case for all profiles P.

In words, R_{RY} consists in duplicating judgment sets in P in a minimal way so that P becomes majority-consistent. R_{RY} is majority-preserving: when P is majority-consistent, no judgment set in P needs to be duplicated to restore majority-consistency.

R_Y, R_{IY} and R_{RY} consider a judgment set as a unit, which is either selected or removed as a whole. Instead of removing entire judgment sets, we may look for finer changes in judgment sets so that the resulting profile becomes majority-consistent. We give two such rules below, defined on the notion of *rectangle* and *co-rectangle* for a profile.

Definition 20 (Rectangles and co-rectangles)**.** *Given a profile $P = \langle J_1, \ldots, J_n \rangle$, we define a* rectangle *for P as a Cartesian product $\rho = I \times Y$, where $I \subseteq \{1, \ldots, n\}$ is a subset of agents and $Y \subseteq \overline{\mathcal{A}}$ is a sub-agenda of $\overline{\mathcal{A}}$. A* co-rectangle *δ for P is the complement of a rectangle for P.*

The restriction of P to rectangle $\rho = I \times Y$ is the profile defined by the set of agents I, the agenda Y, and defined by $P_\rho = \langle J_i \cap Y | i \in I \rangle$.

The restriction of P to co-rectangle $\delta = \overline{I \times Y}$ is the incomplete profile defined by the set of agents $N = \{1, \ldots, n\}$, the agenda $\overline{\mathcal{A}}$, and defined by $P_\delta = \langle J_i \cap Y | i \in N \rangle$.

The intuition for P_ρ is that only the opinions $\langle i, \varphi \rangle$ inside the rectangle count, whereas for P_δ, only the opinions outside the rectangle $\overline{\delta}$ count.

Definition 21 (Maximal rectangle rule R_{MR}). *A rectangle ρ is* maximal P-consistent *if P_ρ is majority-consistent and for every super rectangle ρ' of ρ, $P_{\rho'}$ is not majority-consistent. The* maximal rectangle rule *is defined by*

$$R_{MR}(P) = \{M(P_\rho) \mid \rho \textit{ maximal } P\textit{-consistent}\}$$

Definition 22 (Maximal co-rectangle rule R_{MCR}). *A co-rectangle δ is* maximal P-consistent *if P_ρ is majority-consistent and for every super co-rectangle δ' of δ, $P_{\delta'}$ is not majority-consistent. The* maximal co-rectangle rule *is defined by*

$$R_{MCR}(P) = \{M(P_\delta) \mid \delta \textit{ maximal } P\textit{-consistent}\}$$

R_{MR} are and R_{MCR} are majority-preserving.

Note that if we restrict our attention to (co-)rectangles of P of the form $N \times Y$, then we recover $MSA(P)$, whereas if we restrict our attention to (co-)rectangles of the form $I \times \overline{\mathcal{A}}$, then we recover $mSP(P)$. Similar rules can be obtained by maximizing the *size* of the (co-)rectangle. Before going further, we first establish that R_{MCR} coincides with a rules that we already know.

Proposition 2.2.13. *For all $P \in \Phi^n$, $R_{MSA}(P) = R_{MCR}(P)$.*

Proof. We first prove that, for all $P \in \Phi^n$, if $J \in R_{MCR}(P)$ then $J \in R_{MSA}(P)$.

Let $J \in R_{MCR}(P)$. As a consequence, there exists a maximal P-consistent co-rectangle $\delta = \overline{I \times Y}$ such that $J = M(P_\delta)$. J is consistent, therefore $J \cap M(P)$ is a consistent subset of $M(P)$. Assume that $J \cap M(P) \notin R_{MSA}(P)$. There exists a $J' \in R_{MSA}(P)$ such that $J \cap M(P) \subset J'$. Let $\varphi \in J' \backslash (J \cap M(P))$. From $\varphi \in J' \subseteq M(P)$ and $\varphi \notin J$ we get $\varphi \in Y$. Consider now the co-rectangle $\delta' = \overline{I \times (Y \backslash \{\varphi\})}$. $M(P^\delta)$ and $M(P^{\delta'})$ agree on all elements of the preagenda except φ. Moreover, $M(P^{\delta'})$ and $M(P)$ agree on φ, whereas $M(P^\delta)$ and $M(P)$ disagree on φ. Therefore,

$$M(P^{\delta'} \cap M(P) = (M(P^\delta \cap M(P)) \cup \{\varphi\} \tag{2.1}$$

Lastly, since P^Y is majority-consistent, $P^{\delta'}$ is majority-consistent as well, which together with (2.1), contradicts the maximality of δ. Therefore, $J \in R_{MSA}(P)$.

We now prove that, for all $P \in \Phi^n$, if $J \in R_{MSA}(P)$ then $J \in R_{MCR}(P)$.

Let $Y \in MSA(P)$. As a consequence $M(P) \in R_{MSA}(P)$. Assume that $N \times Y \notin MCR(P)$. Note that $N \times Y = \overline{N \times (\overline{\mathcal{A}} \backslash Y)}$ is a co-rectangle that is P-consistent, therefore there must exists a larger P-consistent co-rectangle. Such a co-rectangle δ' cannot be obtained by removing less agents, since $N \times Y$ does not remove any agent. Consequently we must remove a smaller subset of the agenda, *i.e.*, $\delta' = \overline{N \times (\overline{\mathcal{A}} \backslash Z)} = N \times Z$ with $Z \supset Y$. But then the restriction of P to Z would be majority-consistent, contradicting $Y \in MSA(P)$. Therefore, $Y \in MCR(P)$. □

The last rule we define does not remove agenda elements and/or voters, but looks for a minimal number of *atomic changes* in the profile so that P becomes majority-consistent. We consider an atomic change to be the change of truth value of one element of the preagenda

$Voters$	$p \wedge r$	$p \wedge s$	q	$p \wedge q$	t
$6\times$	$+$	$+$	$+$	$+$	$+$
$4\times$	$+$	$+$	$-$	$-$	$+$
$3\times$	$-$	$-$	$+$	$+$	$-$
$4\times$	$-$	$-$	$+$	$-$	$-$
$M(Q)$	$+$	$+$	$+$	$+$	$+$

Table 2.2: The profile Q.

in an individual judgment set. For instance, if $J_1 = \{p, q, p \wedge q, r, p \wedge r\}$, then $J_1' = \{\neg p,\ q,\ \neg(p \wedge q), r, \neg(p \wedge r)\}$ is obtained from J_1 by a series of three atomic changes (change in the truth value of p, of $p \wedge q$ and of $p \wedge r$).

This approach is in spirit close to Dodgson's voting rule, which looks for the smallest number of elementary changes in a profile with the purpose of turning it into a profile for which a Condorcet winner exists. Replacing *having a Condorcet winner* by *being majority-consistent* and adapting the notion of elementary change, we get our judgment aggregation rule.

Definition 23 (Minimal number of atomic changes rule R_{MNAC})**.** *Given a profile P, a profile Q consisting of complete and consistent individual judgment sets is a closest majority-consistent profile to P if Q is majority-consistent, and there is no majority-consistent profile Q' such that $D_H(P, Q') < D_H(P, Q)$. Let $CMC(P)$ the set of all closest majority-consistent profile to P. The* minimal number of atomic changes rule *is defined by*

$$R_{MNAC}(P) = \{M(Q) \mid Q \in CMC(P)\}$$

R_{MNAC} is not a new rule. $Full_d$, one of the four methods introduced by (Miller and Osherson, 2009), looks for the closest profile of individual judgments that yields a consistent proposition-wise majority output, and then take this output. Therefore, R_{MNAC} corresponds to the $Full_d$ voting rule together with the choice of the Hamming distance as the distance measure d. Miller and Osherson (2009) do not commit to a specific distance metric. Another possible choice would consist in allowing the modified profile to be individually inconsistent, leading to the so-called $Output_d$ rule in (Miller and Osherson, 2009).

Example 2.2.14. *Consider the profile P from Example 2.1.1. The profile Q given on Table 2.2 is the closest majority-consistent profile to P with $D(P, Q) = 3$.*

We obtain $R_{MNAC}(P) = \{p \wedge r, p \wedge s, q, p \wedge q, t\}$.

If P is majority-consistent then no elementary change is needed, therefore R_{MNAC} is majority-preserving.

We could also look for the closest profiles Q with respect to set inclusion. However, this would give a very weak rule R where φ belongs to some judgment set of $R(P)$ as soon as one individual judgment contains φ.

2.2.4 Rules based on distances

Two classes of distance-based rules appear in the judgment aggregation literature. The first one is characterized by the minimization of distances between judgment sets and does not include altering the profile in any way (Pigozzi, 2006; Endriss et al., 2010b; Miller and Osherson, 2009). This class is derived from distance-based merging operators (Konieczny and Pino-Pérez, 2011). The second one is characterized by the minimization of distances among profiles and relies on making minimal changes to the profile (Miller and Osherson, 2009). The rules we consider in this section resort to some kind of minimization of distances among judgment sets without changing the profile.

Miller and Osherson (2009) propose four distance-based rules for judgment aggregation. We have already discussed three of them, namely `Full`, `Output` and `Endpoint`. The fourth one, `Prototype`, is defined as follows.

Definition 24. *$\texttt{Prototype}_d(J_1,\dots,J_n)$ is the set of all judgment sets $J \in \Phi$ such that $\sum_{i=1}^{n} d(J,J_i) \leqslant \sum_{i=1}^{n} d(J',J_i), \forall J' \in \Phi$.*

This rule has also been considered independently in (Endriss et al., 2010b). We propose a larger family of aggregation rule, in the same spirit as (Miller and Osherson, 2009).

Let $d : \Phi \times \Phi \mapsto \mathbb{N}^0$ be a distance function between judgment sets from Φ and $\odot : (\mathbb{N}^0)^n \mapsto \mathbb{N}^0$ be a symmetric, non-decreasing function such that, for every $x,\ y,\ x_1,\ \dots, x_n \in \mathbb{N}^0$, has the following properties: $\odot(x,\dots,x) = x$; $\odot(x_1,\dots,x_n) = 0$ if and only if $x_1 = \dots = x_n = 0$.

The distance-based judgment aggregation rule $R^{d,\odot}$ induced by d and $\odot$ is defined by:

$$R^{d,\odot}(J_1,\dots,J_n) = \underset{J\in\Phi}{\arg\min}\ \odot(d(J,J_1),\dots d(J,J_n)).$$

Definition 25. *A judgment aggregation rule is* distance-based *if it is equal to $R^{d,\odot}$ for some d and $\odot$.*

Here we consider only $\odot = \sum$ and $\odot = \max$, and the Hamming distance d_H. In the case when $\odot = \sum$ we obtain the distance-based procedure of Endriss et al. (2010b). We choose $R^{d_H,\Sigma}$ because it captures the intuition of a majoritarian operator and $R^{d_H,max}$ because it captures the intuition of compromise between the individuals' judgments (Brams et al., 2007; Konieczny and Pino-Pérez, 2011). The minimization of the maximum distance minimizes the disagreement with the least satisfied individual, hence guaranteeing some degree of compromise.

We show that $R^{d_H,\Sigma}$ and R_{MWA} are equal rules.

Proposition 2.2.15. *For all $P \in \Phi^n$, $R^{d_H,\Sigma}(P) = R_{MWA}(P)$.*

Proof. Given two complete judgment sets J and J', and $\varphi \in \overline{\mathcal{A}}$, define $h(\varphi,J,J') = 1$ if $\varphi \in (J \backslash J') \cup (J' \backslash J)$ and $h(\varphi,J,J') = 0$ otherwise.

For any profile $P = \langle J_1,\dots,J_n\rangle \in \Phi^n$ and any $J \in \Phi$, we have

$$
\begin{aligned}
&\textstyle\sum_{i=1}^{n} d_H(J,J_i)\\
&= \textstyle\sum_{i=1}^{n}\sum_{\varphi\in\overline{\mathcal{A}}} h(\varphi,J,J_i)\\
&= \textstyle\sum_{i=1}^{n}\left(\sum_{\varphi\in J} h(\varphi,J,J_i)+\sum_{\varphi\notin J} h(\varphi,J,J_i)\right)\\
&= \textstyle\sum_{i=1}^{n}\left(\sum_{\varphi\in J} h(\varphi,J,J_i)+\sum_{\neg\varphi\in J} h(\varphi,J,J_i)\right)\\
&= \textstyle\sum_{i=1}^{n}\left(\sum_{\varphi\in J} h(\varphi,J,J_i)+\sum_{\varphi\in J} h(\neg\varphi,J,J_i)\right)\\
&= \textstyle\sum_{\varphi\in J}\left(\sum_{i=1}^{n} h(\varphi,J,J_i)+\sum_{i=1}^{n} h(\neg\varphi,J,J_i)\right)\\
&= \textstyle\sum_{\varphi\in J}\left(n-N(P,\varphi)+N(P,\neg\varphi)\right)\\
&= \textstyle\sum_{\varphi\in J} 2(n-N(P,\varphi))\\
&= 2n*|J|-2w_P(J)
\end{aligned}
$$

Therefore, $\sum_{i=1}^{n} d_H(J,J_i)$ is minimum if and only if $J\in MWA(P)$, that is, $w_P(J)$ is maximum. Since every element of $MWA(P)$ is a complete judgment set, $MWA(P)$ is equal to the set of all complete judgment sets minimizing $\sum_{i=1}^{n} d_H(J,J_i)$, which allows us to conclude that $R^{d_H,\Sigma}$ and R_{WMA} are equivalent.

□

Comparing Definition 25 and the definition of $\texttt{Prototype}_d$ we observe that for all profiles $P\in\Phi^n$, $R^{d_H,\Sigma}(P)=\texttt{Prototype}_{d_H}(P)$. Consequently, for all profiles $P\in\Phi^n$, $R_{MWA}(P)=\texttt{Prototype}_{d_H}(P)$.

As a consequence, $R^{d_H,\Sigma}$ is majority-preserving. This is however not the case for $R^{d_H,max}$, which is the only one of our rules failing to satisfy majority-preservation.

Proposition 2.2.16. *$R^{d_H,max}$ is not majority-preserving.*

Proof. Consider the agenda $\overline{\mathcal{A}}=\{a,\neg a,b,\neg b\}$ and $P=\langle\{a,b\},\{a,b\},\{\neg a,\neg b\}\rangle$. Then $R^{d_H,max}(P)=\{\{a,\neg b\},\{\neg a,b\}\}$; however, P is majority-consistent and $M(P)=\{\{a,b\}\}$.

□

Example 2.2.17. *Consider the profile P for agenda $\overline{\mathcal{A}}$ of Example 2.1.1. We obtain that $R^{d_H,\Sigma}=\{\{p\wedge r,p\wedge s,q,p\wedge q,t\}\}$ while*

$$
R^{d_H,max}(P)=\left\{\begin{array}{rrrrr}
\{\neg(p\wedge r), & \neg(p\wedge s), & q, & \neg(p\wedge q), & t\},\\
\{\neg(p\wedge r), & p\wedge s, & \neg q, & \neg(p\wedge q), & t\},\\
\{\neg(p\wedge r), & p\wedge s, & q, & p\wedge q, & t\},\\
\{p\wedge r, & \neg(p\wedge s), & \neg q, & \neg(p\wedge q), & t\},\\
\{p\wedge r, & \neg(p\wedge s), & q, & p\wedge q, & t\},\\
\{p\wedge r, & p\wedge s, & \neg q, & \neg(p\wedge q), & \neg t\},\\
\{p\wedge r, & p\wedge s, & \neg q, & p\wedge q, & \neg t\}
\end{array}\right\}.
$$

The full calculations are presented inwards Table 2.3.

2.3 (Non)inclusion relationships between the rules

In this section we consider the equality and inclusion relationships between the rules we have introduced. This analysis is necessary to establish how the collective judgments derived

J_i	$\{p \wedge r,$	$p \wedge s,$	$q,$	$p \wedge q,$	$t\}$	$d_H(J, J_{20})$	$d_H(J_i, J_{18})$	$d_H(J_i, J_3)$	$\sum$	max
J_1	-	-	-	-	-	5	3	1	49	5
J_2	-	-	-	-	+	4	2	2	46	4
J_3	-	-	+	-	-	4	4	0	40	4
J_4	-	-	+	-	+	3	3	1	37	3
J_5	-	-	+	+	-	3	5	1	45	5
J_6	-	-	+	+	+	2	4	2	42	4
J_7	-	+	-	-	-	4	2	2	46	4
J_8	-	-	+	+	+	2	4	2	42	4
J_9	-	+	-	-	-	4	2	2	46	4
J_{10}	-	+	-	-	+	3	1	3	43	3
J_{11}	-	+	+	+	-	2	4	2	42	4
J_{12}	-	+	+	+	+	1	3	3	39	3
J_{13}	+	-	-	-	-	4	2	2	46	4
J_{14}	+	-	-	-	+	3	1	3	43	3
J_{15}	+	-	+	+	-	2	4	2	42	4
J_{16}	+	-	+	+	+	1	3	3	39	3
J_{17}	+	+	-	-	-	3	1	3	42	3
J_{18}	+	+	-	-	+	2	0	4	40	4
J_{19}	+	+	+	+	-	1	3	3	39	3
J_{20}	+	+	+	+	+	0	2	4	36	4

Table 2.3: The calculations for $R^{d_H,\Sigma}(P)$ and $R^{d_H,max}(P)$. Recall that $\sum_{i=j}^{n} d(J_i, J_j) = 6d_H(J_i, J_{20}) + 4d_H(J_i, J_{18}) + 7d_H(J_i, J_3)$

from one rule compare to the collective judgments derived from another rule. We have the following diagram (Table 2.4), where *inc* means "inclusion-wise incomparable", $\subset$ means that $T_{R_1}(P) \subset T_{R_2}(P)$ for every profile $P \in \Phi^n$, where R_1 is the row rule and R_2 is the column rule, correspondingly for $\supset$. The number next to inc, $\subset$ or $\supset$, denotes the proposition in which the relationship is proved.

	R_{MCSA}	R_{MWA}	R_{RA}	R_Y	R_{IY}	R_{RY}	R_{MR}	R_{MNAC}	$R^{d_H,\max}$
R_{MSA}	$\subset$, 2.3.1	$\subset$, 2.3.2	$\subset$, 2.3.3	inc,2.3.5	inc,2.3.16,2.3.21	inc,2.3.14	$\supset$,2.3.18	inc, 2.3.22	inc, 2.3.4
R_{MCSA}		inc,2.3.6	inc,2.3.7	inc,2.3.5	inc, 2.3.16,2.3.21	inc,2.3.12	$\supset$,2.3.19	inc,2.3.21	inc,2.3.4
R_{MWA}			inc,2.3.10	inc,2.3.8	inc, 2.3.21,2.3.16	inc,2.3.12	$\supset$,2.3.19	inc,2.3.23	inc,2.3.4
R_{RA}				inc,2.3.9	inc, 2.3.16,2.3.21	inc,2.3.12	$\supset$, 2.3.20	inc,2.3.23	inc,2.3.4
R_Y					$\supset$,2.3.15	inc,2.3.11	$\supset$,2.3.19	inc,2.3.23	inc,2.3.4
R_{IY}						inc 2.3.11	$\supset$,2.3.18	inc,2.3.23	inc,2.3.4
R_{RY}							inc2.3.13	inc,2.3.23	inc,2.3.4
R_{MR}								inc, 2.3.24	inc,2.3.4
R_{MNAC}									inc,2.3.4

Table 2.4: A summary of the (non)inclusion relationships between the proposed rules.

For every profile $P \in \Phi^n$, if a collective judgment is in all the judgment sets $R_{MCSA}(P)$, or $R_{MWA}(P)$ or $R_{RA}(P)$, then that collective judgment is in all the judgment sets $R_{MSA}(P)$. For ever profile $P \in \Phi^n$, if a collective judgment is in all the judgment sets $R_{MSA}(P)$, $R_{MCSA}(P)$, $R_{MWA}(P)$, $R_{RA}(P)$, $R_Y(P)$ or $R_{IY}(P)$, then that collective judgment is in all the judgment sets $R_{MR}(P)$. This means that the rules R_{MSA} and R_{MR} are very "weak" in the sense that they often select a very large number of judgment sets. In this sense the rule R_{IY} is weaker than R_Y. For a decision reaching context in which the rule should select as little judgment sets as possible, we can choose from seven rules, out of which five, R_{MCSA}, R_{RA}, R_Y, R_{RY} and $R^{d_H,max}$ have not been studied in judgment aggregation.

Following are the proofs for these (non) inclusion relationships.

Proposition 2.3.1. *For every profile* $P \in \Phi^n$, $T_{R_{MSA}}(P) \subset T_{R_{MCSA}}(P)$.

Proof. If $Y \subset \mathcal{A}$ is a maxcard consistent sub-preagenda of $\mathcal{A}$, with respect to P, then it is also a maximal consistent sub-preagenda with respect to P. If $\alpha \in T_{R_{MSA}}(P)$, then α is inferred in every maximal consistent sub-preagenda, and *a fortiori* in every maxcard consistent sub-preagenda, therefore $\alpha \in T_{R_{MCSA}}(P)$.

To show that $T_{R_{MCSA}}(P) \nsubseteq T_{R_{MSA}}(P)$, consider the profile P in Example 2.1.1. As it can be observed in Example 2.2.2, $T_{R_{MCSA}}(P) \models p \wedge r$, but we can observe from Example 2.2.1 that $T_{R_{MSA}}(P) \not\models p \wedge r$. □

Proposition 2.3.2. *For every profile* $P \in \Phi^n$, $T_{R_{MSA}}(P) \subset T_{R_{MWA}}(P)$.

Proof. If $Y \subset \mathcal{A}$ is a consistent sub-preagenda maximizing $w_P(Y)$, then $M(P^{\downarrow Y})$ is a maximal consistent sub-agenda with respect to P. If $\alpha \in T_{R_{MSA}}(P)$, then α is inferred in every maximal consistent sub-preagenda, and *a fortiori* in every maxweight consistent sub-agenda, therefore $\alpha \in T_{R_{MWA}}(P)$.

To show that $T_{R_{MWA}}(P) \nsubseteq T_{R_{MSA}}(P)$, consider the profile P in Example 2.1.1. As it can be observed in Example 2.2.6, $T_{R_{MWA}}(P) \models q$, but we can observe from Example 2.2.1 that $T_{R_{MSA}}(P) \not\models q$.

□

Proposition 2.3.3. *For every profile* $P \in \Phi^n$, $T_{R_{MSA}}(P) \subset T_{R_{RA}}(P)$.

Proof. In the construction of $R_{RA}(P)$, let Z be the subset of $\overline{\mathcal{A}}$ composed of the ψ_k such that $\delta \wedge \psi_k$ is consistent. Z is a maximal consistent sub-agenda with respect to P. Z is consistent by construction, and maximal because every time a formula ψ_k is rejected, it is because it produces an inconsistency with the formulas already present in δ. If $\alpha \in T_{R_{MSA}}(P)$, then α is inferred in every maximal consistent sub-agenda, and *a fortiori* in Z, therefore $\alpha \in T_{R_{RA}}(P)$.

To show that $T_{R_{RA}}(P) \nsubseteq T_{R_{MSA}}(P)$, consider the profile P in Example 2.1.1. As it can be observed in Example 2.2.8, $T_{R_{RA}}(P) \models q$, but we can observe from Example 2.2.1 that $T_{R_{MSA}}(P) \not\models q$. □

Proposition 2.3.4. $R^{d_H,\max}$ *is incomparable with all the other rules.*

Proof. Let R be a majority-preserving rule. Take the profile P as in the proof of Proposition 2.2.16. Then $a \leftrightarrow \neg b \in T_{R^{d_H,\max}}(P)$, whereas $a \leftrightarrow \neg b \notin T_R(P)$ (since $a \leftrightarrow b \in T_R(P)$); and $a \in T_R(P)$, whereas $a \notin T_{R^{d_H,\max}}(P)$. Therefore, $R^{d_H,\max}$ is incomparable with all of the five other rules. □

Proposition 2.3.5. R_Y *is incomparable with* R_{MSA} *and* R_{MCSA}.

a	$a \to (b \vee c)$	b	c	$a \to (d \vee e)$	d	e
+	+	+	−	+	+	−
+	+	−	+	+	−	+
+	−	−	−	−	−	−

Table 2.5: The profile P used in proving R_Y *inc* R_{MSA} and R_Y *inc* R_{MCSA}.

Proof. Consider the pre-agenda $\mathcal{A} = \{a, a \to (b \vee c), b, c, a \to (d \vee e), d, e\}$ and the profile P for this agenda for three agents given on Table 2.5.

We have that $M(P) = \{a, a \to (b \vee c), \neg b, \neg c, a \to (d \vee e), \neg d, \neg e\}$. The minimal inconsistent subsets of $M(P)$ are $\{a, a \to (b \vee c), b, c\}$ and $\{a, a \to (d \vee e), d, e\}$. Consequently $M(P)$ has 10 maximal consistent subsets: 9 containing a and one equal to $M(P) \setminus \{a\}$. The 9 sets containing a contain two of the three formulas $\{a \to (b \vee c), \neg b, \neg c\}$ and two of the three formulas $\{a \to (d \vee e), \neg d, \neg e\}$. These 10 maximal consistent subsets correspond to 10 maximal sub-agendas. The only maxcard consistent sub-agenda is $\mathcal{A} \setminus \{a\}$, and in this sub-agenda of $\mathcal{A}$, $\neg a$ is inferred. Therefore, $T_{R_{MCSA}}(P) \models \neg a$. All sub-profiles of P of size two are majority-consistent, and each of them accepts a, therefore $T_{R_Y}(P) \models a$. As a consequence, R_Y and R_{MCSA} are incomparable. For $T_{R_Y}(P) \nsubseteq T_{R_{MSA}}(P)$, take the same profile as above and note that $a \in T_{R_Y}(P)$ but $a \notin T_{R_{MSA}}(P)$.

To show that $T_{R_{MSA}}(P) \nsubseteq T_{R_Y}(P)$, assume the pre-agenda is extended with another agenda item φ, on which the agents vote +, + and - correspondingly. We have $\varphi \in T_{R_{MSA}}$ but $\varphi \notin T_{R_Y}$. □

Proposition 2.3.6. R_{MWA} is incomparable with R_{MCSA}.

Proof. To show that $R_{MWA} \not\subset R_{MCSA}$ take the following seven agent profile P:

	a	b	$a \wedge b$
3×	+	+	+
2×	+	−	−
2×	−	+	−

We obtain $R_{MWA}(P) = \{\{a, b, a \wedge b\}\}$ and $R_{MCSA}(P) = \{\{a, b\}, \{a, \neg a \vee \neg b\}, \{b, \neg a \vee \neg b\}\}$. Consequently, $a \in T_{R_{MWA}}(P)$ and $a \notin T_{R_{MCSA}}(P)$.

For the converse, that $R_{MCSA} \not\subset R_{MWA}$ revisit the example of Proposition 2.3.5. We have $\neg a \notin T_{R_{MWA}}(P)$ and $\neg a \in T_{R_{MCSA}}(P)$. □

Proposition 2.3.7. *R_{RA} is incomparable with R_{MCSA}.*

Proof. Same profile P as in Proposition 2.3.5. We have that $T_{R_{RA}}(P) \models a$. Hence $a \in T_{R_{RA}}(P)$ and $\neg a \in R_{MCSA}(P)$, see Proposition 2.3.5. □

Proposition 2.3.8. *R_{MWA} is incomparable with R_Y.*

Proof. Consider the following pre-agenda
$\mathcal{A} = \{a, a \to p_1, a \to q_1, a \to (p_1 \wedge q_1), a \to p_2, a \to q_2, a \to (p_2 \wedge q_2), a \to p_3,$
$a \to q_3, a \to (p_3 \wedge q_3), a \to p_4, a \to q_4, a \to (p_4 \wedge q_4)\}$.

Let the profile P be as given on Table 2.6.

Agenda	Voters ×1	×1	×1	$M(P)$	$N(P,\varphi_i)$
a	+	+	+	+	3
$a \to p_1$	+	+	-	+	2
$a \to q_1$	+	-	+	+	2
$a \to (p_1 \wedge q_1)$	+	-	-	-	1
$a \to p_2$	+	+	-	+	2
$a \to q_2$	+	-	+	+	2
$a \to (p_2 \wedge q_2)$	+	-	-	-	1
$a \to p_3$	+	+	-	+	2
$a \to q_3$	+	-	+	+	2
$a \to (p_3 \wedge q_3)$	+	-	-	-	1
$a \to p_4$	+	+	-	+	2
$a \to q_4$	+	-	+	+	2
$a \to (p_4 \wedge q_4)$	+	-	-	-	1

Table 2.6: The profile P used to show R_{MWA} *inc* R_Y. The judgment sets are the second, third and fourth column of the table.

We obtain that $R_{MWA}(P) = \{\{\neg a, a \to p_1, a \to q_1, \neg(a \to (p_1 \wedge q_1)), a \to p_2, a \to q_2, \neg(a \to (p_2 \wedge q_2)), a \to p_3, a \to q_3, \neg(a \to (p_3 \wedge q_3)), a \to p_4, a \to q_4, \neg(a \to (p_4 \wedge q_4))\}\}$. Hence $T_{R_{MWA}}(P) \models \neg a$.

The result for $R_Y(P)$ is obtained when exactly one, either one, of the voters is removed. For $R_Y(P)$ we obtain $T_{R_Y}(P) \models a$. ☐

Proposition 2.3.9. *R_{RA} is incomparable with R_Y.*

Proof. We have $T_{R_{RA}}(P) \nsubseteq T_{R_Y}(P)$ as a consequence of Propositions 2.3.3 and 2.3.5.

To show that $T_{R_Y}(P) \nsubseteq T_{R_{RA}}(P)$, consider the pre-agenda $\mathcal{A} = \{p, q, p \wedge q, r, s, r \wedge s, t\}$ and the 18 agents profile P:

	p	q	$p \wedge q$	r	s	$r \wedge s$	t
1×	+	+	+	−	+	−	+
3×	+	+	+	−	+	−	−
4×	+	+	+	+	−	−	−
2×	+	−	−	+	−	−	−
4×	+	−	−	+	+	+	+
4×	−	+	−	+	+	+	+

The minimal number of agents to remove to make the profile majority-consistent is two. These two agents are the two agents of the fourth row (light gray shaded). We obtain $t \in T_{R_Y}(P)$ and $t \notin T_{R_{RA}}(P)$. ☐

Proposition 2.3.10. *$R_{RA}(P)$ is incomparable with R_{MWA}.*

Proof. Consider the same profile example in Proposition 2.3.8. We obtain that $\varphi_{13} \in T_{R_{RA}}(P)$ and $\neg\varphi_{13} \in T_{R_{MWA}}(P)$.

☐

Proposition 2.3.11. *R_{RY} is incomparable with R_Y and R_{IY}.*

Proof. Consider the pre-agenda $\mathcal{A} = \{p, q, p \wedge q, r, s, r \wedge s, t\}$ and the profile P for this agenda given in Table 2.7.

Voters	p	q	$p \wedge q$	r	s	$r \wedge s$	t
1×	+	+	+	−	+	−	+
3×	+	+	+	−	+	−	−
4×	+	+	+	+	−	−	−
2×	+	−	−	+	−	−	−
4×	+	−	−	+	+	+	+
4×	−	+	−	+	+	+	+

Table 2.7: The profile P used to show $T_{R_{RY}}(P) \models \neg t$.

We obtain $R_{RY}(P) = \{p, q, p \wedge q, r, s, r \wedge s, \neg t\}$, by adding the fourth judgment set six times, *i.e.*, as $M(P')$, where P' is the profile given in Table 2.8. Consequently $T_{R_{RY}}(P) \models \neg t$.

Voters	p	q	$p \wedge q$	r	s	$r \wedge s$	t
1×	+	+	+	−	+	−	+
3×	+	+	+	−	+	−	−
4×	+	+	+	+	−	−	−
8×	+	−	−	+	−	−	−
4×	+	−	−	+	+	+	+
4×	−	+	−	+	+	+	+
$M(P')$	+		−	+		−	−

Table 2.8: The profile P' obtained from P of Table 2.7 by adding the fourth judgment set 4 times.

The $R_Y(P)$ and $R_{IY}(P)$ are obtained by calculating $M(P'')$, the profile P'' being given in Table 2.9 and obtained from P by removing the fourth judgment set (both of them).
$R_Y(P) = R_{IY}(P) = \{p, q, p \wedge q, r, s, r \wedge s, t\}$, hence $T_{R_Y}(P) \models t$ and $T_{R_{IY}} \models t$.

Voters	p	q	$p \wedge q$	r	s	$r \wedge s$	t
1×	+	+	+	−	+	−	+
3×	+	+	+	−	+	−	−
4×	+	+	+	+	−	−	−
0×	+	−	−	+	−	−	−
4×	+	−	−	+	+	+	+
4×	−	+	−	+	+	+	+
$M(P'')$	+	+	+	+	+	+	+

Table 2.9: The profile P'' obtained from P of Table 2.7 by removing the two judgment sets in the fourth row.

□

Proposition 2.3.12. *R_{RY} is incomparable with R_{MCSA} and R_{MWA}.*

Proof. Consider the pre-agenda from Example 2.1.1 and the profile P given in Table 2.1. As it can be observed in Example 2.2.12, $R_{RY}(P) = \{\neg(p \wedge r), \neg(p \wedge s), q, \neg(p \wedge q), t\}$, obtained by adding 3 of the judgment sets $\{\neg(p \wedge r), \neg(p \wedge s), q, \neg(p \wedge q), \neg t\}$.
Hence $T_{R_{RY}}(P) \models \neg(p \wedge r)$.

For the same profile P we obtain that $T_{R_{MCSA}}(P) \models p \wedge r$ and $T_{R_{MWA}}(P) \models p \wedge r$, see Examples 2.2.2 and 2.2.6. □

Proposition 2.3.13. *R_{RY} is incomparable with R_{MR}.*

Voters	p	q	$p \wedge q$	r	s	$r \wedge s$	t	u	v	w	y
$1 \times J_1$	+	+	+	+	+	+	−	−	−	−	−
$1 \times J_2$	+	+	+	+	−	−	+	+	−	−	+
$1 \times J_3$	+	+	+	+	−	−	+	+	−	+	−
$1 \times J_4$	+	−	−	+	+	+	+	−	+	−	−
$1 \times J_5$	+	−	−	+	+	+	−	+	+	+	+
$1 \times J_6$	−	+	−	−	+	−	+	−	+	+	+
$1 \times J_7$	−	+	−	−	+	−	−	+	+	+	+
Majority	+	+	−	+	+	−	+	+	+	+	+

Table 2.10: The profile P used to prove R_{MR} *inc* R_{RY}.

Proof. Consider the pre-agenda $\mathcal{A} = \{p, q, p \wedge q, r, s, r \wedge s, t, u, v, w, y\}$ and the Profile P for it given on Table 2.10. We obtain $R_{RY}(P)$, by repeating the first judgment set once, namely, as $M(P_{RY})$, P_{RY} being given on Table 2.11. We obtain that $R_{RY}(P) = \{p, q, p \wedge q, r, s, r \wedge s\}$ and as consequence $T_{R_{RY}}(P) \not\models t \vee u \vee v \vee w \vee y$.

Voters	p	q	$p \wedge q$	r	s	$r \wedge s$	t	u	v	w	y
$2 \times J_1$	+	+	+	+	+	+	−	−	−	−	−
$1 \times J_2$	+	+	+	+	−	−	+	+	−	−	+
$1 \times J_3$	+	+	+	+	−	−	+	+	−	+	−
$1 \times J_4$	+	−	−	+	+	+	+	−	+	−	−
$1 \times J_5$	+	−	−	+	+	+	−	+	+	+	+
$1 \times J_6$	−	+	−	−	+	−	+	−	+	+	+
$1 \times J_7$	−	+	−	−	+	−	−	+	+	+	+
Majority	+	+	−	+	+	−	+	+	+	+	+

Table 2.11: The profile P_{RY} constructed from profile P in Table 2.10.

The maximal rectangles are given in Table 2.12. We obtain, $T_{R_{MR}}(P) \models t \vee u \vee v \vee w \vee y$. Consequently $T_{R_{MR}}(P) \not\subseteq T_{R_{RY}}(P)$.

Now consider the pre-agenda $\mathcal{A} = \{p, q, p \wedge q\}$ and for it the profile P in Table 2.13.

We obtain $R_{RY}(P) = \{\{p, q, p \wedge q\}, \{p, \neg q, \neg(p \wedge q)\}, \{\neg p, q, \neg(p \wedge q)\}\}$. Consequently, $T_{R_{RY}}(P) \models p \vee q$. The maximal rectangles for this profile are:

Set	Removed sets	Removed issues	*ResultingSets*
J_8	$\{J_6\}$		$\{p,\ q,\ r,\ s,\ u\}$
J_9	$\{J_7\}$		$\{p,\ q,\ r,\ s,\ t\}$
J_{10}	$\{J_1,J_2,J_3\}$		$\{\neg(p \wedge q),\ s,\ v,\ w,\ y\}$
J_{11}	$\{J_1,J_2,J_4\}$		$\{q,\ \neg(p \wedge q),\ s,\ \neg(r \wedge s),\ u,\ v,\ w,\ y\}$
J_{12}	$\{J_1,J_2,J_5\}$		$\{q,\ \neg(p \wedge q),\ s,\ \neg(r \wedge s),\ t,\ v,\ w\}$
J_{13}	$\{J_1,J_3,J_4\}$		$\{q,\ \neg(p \wedge q),\ s,\ \neg(r \wedge s),\ u,\ v,\ w,\ y\}$
J_{14}	$\{J_1,J_3,J_5\}$		$\{q,\ \neg(p \wedge q),\ s,\ \neg(r \wedge s),\ t,\ v,\ y\}$
J_{15}	$\{J_1,J_4,J_5\}$		$\{q,\ \neg(r \wedge s),\ t,\ u,\ v\ w,\ y\}$
J_{16}	$\{J_2,J_3,J_4\}$		$\{q,\ \neg(p \wedge q)\ s,\ \neg t\ v\ w\ y\}$
J_{17}	$\{J_2,J_4,J_5\}$		$\{q,\ s,\ \neg(r \wedge s),\ w\}$
J_{18}	$\{J_2,J_3,J_5\}$		$\{p,\ q,\ s,\ \neg(r \wedge s),\ \neg u,\ v\}$
J_{19}	$\{J_3,J_4,J_5\}$		$\{q,\ n\ s,\ \neg(r \wedge s),\ y\}$
J_{20}		$\{p,r\}$	$\{q,\ s,\ \neg(r \wedge s),\ t,\ u,\ v,\ w,\ y\}$
J_{21}		$\{p,s\}$	$\{q,\ r,\ \neg(r \wedge s),\ t,\ u,\ v,\ w,\ y\}$
J_{22}		$\{p,r \wedge s\}$	$\{q,\ r,\ s,\ t,\ u,\ v,\ w,\ y\}$
J_{23}		$\{q,r\}$	$\{p,\ q,\ s,\ \neg(r \wedge s),\ t,\ u,\ v,\ w,\ y\}$
J_{24}		$\{q,s\}$	$\{p,\ q,\ r,\ \neg(r \wedge s),\ t,\ u,\ v,\ w,\ y\}$
J_{25}		$\{q,r \wedge s\}$	$\{p,\ q,\ r,\ s,\ t,\ u,\ v,\ w,\ y\}$
J_{26}		$\{p \wedge q,r\}$	$\{p,\ q,\ s,\ \neg(r \wedge s),\ t,\ u,\ v,\ w,\ y\}$
J_{27}		$\{p \wedge q,s\}$	$\{p,\ q,\ r,\ \neg(r \wedge s),\ t,\ u,\ v,\ w,\ y\}$
J_{28}		$\{p \wedge q,r \wedge s\}$	$\{p,\ q,\ r,\ s,\ t,\ u,\ v,\ w,\ y\}$
J_{29}	$\{J_2\}$	$\{p\}$	$\{q,\ \neg(p \wedge q)\ r,\ s,\ t,\ v,\ w\}$
J_{30}	$\{J_3\}$	$\{p\}$	$\{q,\ \neg(p \wedge q)\ r,\ s,\ t,\ v,\ y\}$
J_{31}	$\{J_2\}$	$\{q\}$	$\{p,\ \neg(p \wedge q)\ r,\ s,\ t,\ v,\ w\}$
J_{32}	$\{J_3\}$	$\{q\}$	$\{p,\ \neg(p \wedge q)\ r,\ s,\ t,\ v,\ y\}$
J_{33}	$\{J_2\}$	$\{p \wedge q\}$	$\{p,\ q,\ r,\ s,\ t,\ v,\ w\}$
J_{34}	$\{J_3\}$	$\{p \wedge q\}$	$\{p,\ q,\ r,\ s,\ t,\ v,\ y\}$
J_{35}	$\{J_4\}$	$\{r\}$	$\{p,\ q,\ s,\ \neg(r \wedge s)\ u,\ w,\ y\}$
J_{36}	$\{J_5\}$	$\{r\}$	$\{p,\ q,\ s,\ \neg(r \wedge s)\ t\}$
J_{37}	$\{J_4\}$	$\{s\}$	$\{p,\ q,\ r,\ \neg(r \wedge s)\ u,\ w,\ y\}$
J_{38}	$\{J_5\}$	$\{s\}$	$\{p,\ q,\ r,\ \neg(r \wedge s)\ t\}$
J_{39}	$\{J_4\}$	$\{r \wedge s\}$	$\{p,\ q,\ r,\ s,\ u,\ w,\ y\}$
J_{40}	$\{J_5\}$	$\{r \wedge s\}$	$\{p,\ q,\ r,\ s,\ t\}$

Table 2.12: Maximal rectangles for the profile in Table 2.10.

Voters	p	q	$p \wedge q$
$1 \times J_1$	+	+	+
$1 \times J_2$	+	−	−
$1 \times J_3$	−	+	−
$M(P)$	+	+	−

Table 2.13: The profile P, counter-example for $R_{RY}(P) \in R_{MR}(P)$.

- $\rho_1 = \{J_2, J_3\} \times \mathcal{A}$ giving rise to $\{\neg(p \wedge q)\}$,
- $\rho_2 = \{J_1, J_3\} \times \mathcal{A}$ giving rise to $\{q\}$,
- $\rho_3 = \{J_1, J_2\} \times \mathcal{A}$ giving rise to $\{p\}$,
- $\rho_4 = N \times \{q, p \wedge q\}$ giving rise to $\{q, \neg(p \wedge q)\}$,
- $\rho_5 = N \times \{p, p \wedge q\}$ giving rise to $\{p, \neg(p \wedge q)\}$,
- $\rho_6 = N \times \{p, q\}$ giving rise to $\{p, q\}$,

Due to $\{\neg(p \wedge q)\} \in R_{MR}(P)$ we obtain that $T_{R_{MR}}(P) \not\models p \vee q$. Hence, $T_{R_{RY}}(P) \nsubseteq T_{R_{MR}}(P)$. □

Proposition 2.3.14. *R_{RY} is incomparable with R_{MSA}.*

Proof. To prove that $T_{R_{MSA}}(P) \nsubseteq T_{R_{RY}}(P)$ we consider the profile P from Example 2.1.1, and Examples 2.2.12 and 2.2.1. We obtain that $T_{R_{MSA}}(P) \models t$, but $T_{R_{RY}}(P) \not\models t$.

To prove that $T_{R_{RY}}(P) \nsubseteq T_{R_{MSA}}(P)$ we consider the profile P from Table 2.7. We obtain $T_{R_{RY}}(P) \models \neg t$. However $T_{R_{MSA}}(P) \not\models \neg t$.

□

Proposition 2.3.15. *For every profile $P \in \Phi^n$, $T_{R_{IY}}(P) \subset T_{R_Y}(P)$.*

Proof. If $Y \subset \overline{\mathcal{A}}$ is a maxcard consistent sub-preagenda (w.r.t. P) of $\overline{\mathcal{A}}$ then it is also a maximal consistent sub-preagenda with respect to P. If $\alpha \in T_{R_{IY}}(P)$, then α is inferred in every maximal consistent sub-agenda and *a fortiori* in every maxcard consistent sub-agenda. Conseqeunetly $\alpha \in T_{R_Y}(P)$.

To show that $T_{R_Y}(P) \not\subset T_{R_{IY}}(P)$ consider the profile from Example 2.1.1. For this profile we obtain:

- $R_Y(P) = \{\{\neg(p \wedge r), \neg(p \wedge s), q, \neg(p \wedge q)\}\}$, see Example 2.2.10;
- $R_{IY}(P) = \{\{\neg(p \wedge r), \neg(p \wedge s), q, \neg(p \wedge q)\}, \{p \wedge r, p \wedge s, q, p \wedge q, t\}\}$, see Example 2.2.11.

It holds that $T_{R_Y}(P) \models \neg(p \wedge r)$, while $T_{R_{IY}}(P) \not\models \neg(p \wedge r)$. □

Proposition 2.3.16. *There exists a $P \in \Phi^n$ such that $T_{R_{IY}}(P) \nsubseteq T_{R_Z}(P)$ for $Z \in \{MSA, MCSA, MWA, RA, RY, MR, MCR, MNAC\}$.*

Proof. Consider the pre-agenda $\mathcal{A} = \{p, q, p \wedge q, t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8\}$ and for it the profile P given on Table 2.14. We obtain R_{IY} by removing any two voters:

Voters	p	q	$p \wedge q$	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9
$J_1 \times$	+	+	+	+	-	-	-	+	+	+	-	-
$J_2 \times$	+	+	+	+	+	-	-	-	+	-	+	-
$J_3 \times$	+	-	-	+	+	+	-	-	-	+	-	+
$J_4 \times$	+	-	-	-	+	+	+	-	-	+	-	-
$J_5 \times$	-	+	-	-	-	+	+	+	-	-	+	+
$J_6 \times$	-	+	-	-	-	-	+	+	+	-	+	+
$M(P)$	+	+	-									

Table 2.14: The profile P.

- removing J_1 and J_2 we obtain $\{\neg(p \wedge q), \neg t_1, t_3, t_4, \neg t_6, t_9\}$
- removing J_1 and J_3 we obtain $\{q, \neg(p \wedge q), \neg t_1, t_4, \neg t_7, t_8\}$

- removing J_1 and J_4 we obtain $\{q, \neg(p \land q), \neg t_7, t_8, t_9\}$
- removing J_1 and J_5 we obtain $\{p, \neg(p \land q), t_2, \neg t_5\}$
- removing J_1 and J_6 we obtain $\{p, \neg(p \land q), t_2, t_3, \neg t_5, \neg t_6\}$
- removing J_2 and J_3 we obtain $\{q, \neg(p \land q), \neg t_1, \neg t_2, t_4, t_5\}$
- removing J_2 and J_4 we obtain $\{q, \neg(p \land q), \neg t_2, t_5, t_9\}$
- removing J_2 and J_5 we obtain $\{p, \neg(p \land q), t_7, \neg t_8\}$
- removing J_2 and J_6 we obtain $\{p, \neg(p \land q), t_3, \neg t_6, t_7, \neg t_8\}$
- removing J_3 and J_4 we obtain $\{q, \neg t_2 \neg t_3, t_5, t_6, \neg t_7, t_8\}$
- removing J_3 and J_5 we obtain $\{p, q, \neg t_3, t_6, \neg t_9\}$
- removing J_3 and J_6 we obtain $\{p, q, \neg t_9\}$
- removing J_4 and J_5 we obtain $\{p, q, t_1, \neg t_3, \neg t_4, t_6\}$
- removing J_4 and J_6 we obtain $\{p, q, t_1, \neg t_4\}$
- removing J_5 and J_6 we obtain $\{p, t_1, t_2, \neg t_4, \neg t_5, t_7 \neg t_8, \neg t_9\}$

We have that $R_{IY}(P) = R_Y(P)$. Let us denote with α the formula $\neg t_1 \lor \neg t_2 \lor \neg t_3 \lor \neg t_4 \lor \neg t_5 \lor \neg t_6 \lor \neg t_7 \lor \neg t_8 \lor \neg t_9$ and with β the formula $t_1 \lor t_2 \lor t_3 \lor t_4 \lor t_5 \lor t_6 \lor t_7 \lor t_8 \lor t_9$. We obtain that $\alpha \in T_{R_{IY}}(P)$, but $\beta \notin T_{R_{IY}}(P)$. For the rest of the rules we obtain:

- $\{\{p\}, \{q\}, \{p \land q\}\} \subset R_Z(P)$, where $Z \in \{MSA, MCSA, MR, MCR\}$ hence $\alpha \notin T_{R_Z}(P)$.
- Observe that $N(P, p) = 4$, $N(P, \neg p) = 2$, $N(P, q) = 4$, $N(P, \neg q) = 2$, $N(P, p \land q) = 2$, $N(P, \neg(p \land q)) = 4$, $N(P, t_i) = N(P, \neg t_i) = 3$, for $i \in [1, 9]$. Observe that a judgment set that include either one of $\{p, q, p\}$, $\{\neg p, q, \neg(p \land q)$ or $\{p, \neg q, \neg(p \land q)$ and any consistent subset of $\{t_i, \neg t_i\}$, $i \in [1, 9]$ will have the maximum weight of 37. Consequently $R_{MWA}(P)$ contains the judgment sets:
  $\{p, q, p \land q, t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8, t_9\}$,
  $\{\neg p, q, \neg(p \land q), t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8, t_9\}$,
  $\{p, \neg q, \neg(p \land q), t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8, t_9\}$.

  Consequently $T_{MWA}(P) \not\models \alpha$.
- For the profile P we can construct the following order:
  $p \sim q \sim \neg(p \land q) > t_1 \sim \neg t_1 \sim t_2 \sim \neg t_2 \sim t_3 \sim \neg t_3 \sim t_4 \sim \neg t_4 \sim t_5 \sim \neg t_5 \sim t_6 \sim \neg t_6 \sim t_7 \sim \neg t_7 \sim t_8 \sim \neg t_8 \sim t_9 \sim \neg t_9 > \neg p \sim \neg q \sim p \land q$.
  Correspondingly, we obtain, among others, the following judgment sets in $R_{RA}(P)$:
  $\{p, q, t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8, t_9\}$,
  $\{p, \neg(p \land q), t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8, t_9\}$,
  $\{q, \neg(p \land q), t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8, t_9\}$,
  $\{p, q, \neg t_1, \neg t_2, \neg t_3, \neg t_4, \neg t_5, \neg t_6, \neg t_7, \neg t_8, \neg t_9\}$,
  $\{p, \neg t_1, \neg t_2, \neg t_3, \neg t_4, \neg t_5, \neg t_6, \neg t_7, \neg t_8, \neg t_9\}$,
  $\{q, \neg t_1, \neg t_2, \neg t_3, \neg t_4, \neg t_5, \neg t_6, \neg t_7, \neg t_8, \neg t_9\}$,
  $\{p, q, t_1, \neg t_2, \neg t_3, \neg t_4, \neg t_5, \neg t_6, \neg t_7, \neg t_8, \neg t_9\}$,

$\{p, t_1, \neg t_2, \neg t_3, \neg t_4, \neg t_5, \neg t_6, \neg t_7, \neg t_8, \neg t_9\}$,
$\{q, t_1, \neg t_2, \neg t_3, \neg t_4, \neg t_5, \neg t_6, \neg t_7, \neg t_8, \neg t_9\}$,
etc.

Consequently $\alpha \notin T_{R_{RA}}(P)$.

- $R_{MNAC} = \{\{p, \neg(p \wedge q)\}, \{q, \neg(p \wedge q)\}\}$, hence $\alpha \notin T_{R_{MNAC}}(P)$.

□

Proposition 2.3.17. *There exists a profile P such that:*

1. $T_{R_{MSA}}(P) \nsubseteq T_{R_{IY}}(P)$,
2. $T_{R_{MCSA}}(P) \nsubseteq T_{R_{IY}}(P)$,
3. $T_{R_{MWA}}(P) \nsubseteq T_{R_{IY}}(P)$,
4. $T_{R_{RA}}(P) \nsubseteq T_{R_{IY}}(P)$.

Proof. This relationship follows from $T_{R_{IY}}(P) \subset T_{R_Y}(P)$, Proposition 2.3.15, and: Proposition 2.3.5, for items 1 and 2, Proposition 2.3.8 for item 3, and Proposition 2.3.9 for item 4. □

Proposition 2.3.18. *For every profile $P \in \Phi^n$, $T_{R_{MR}}(P) \subset T_{R_{MSA}}(P)$ and $T_{R_{MR}}(P) \subset T_{R_{IY}}(P)$.*

Proof. Removing only voters, as we do in R_{IY}, corresponds to a rectangle in the form of $I \times \overline{\mathcal{A}}$, while removing only votes on a subset of the agenda, like we do in R_{MSA}, corresponds to a rectangle in the form of $N \times Y$.

To show that $T_{R_{MSA}} \not\subset (P)T_{R_{MR}}(P)$, consider the profile P from Example 2.1.1. For this profile, we obtain:

- $R_{MSA}(P) = \{\{p \wedge r, p \wedge s, q, t\}, \{p \wedge r, p \wedge s, \neg(p \wedge q), t\}, \{q, \neg(p \wedge q), t\}\}$, see Example 2.2.1; and
- $R_{IY}(P) = \{\{\neg(p \wedge r), \neg(p \wedge s), q, \neg(p \wedge q)\}, \{p \wedge r, p \wedge s, q, p \wedge q, t\}\}$, see Example 2.2.11.

We obtain $T_{R_{MSA}} \models t$, $T_{R_{MSA}} \not\models q$, $T_{R_{IY}} \models q$ and $T_{R_{MSA}} \not\models t$. Since $R_{MSA}(P) \cup R_{IY}(P) \subset R_{MR}(P)$ we obtain that $T_{R_{MR}} \not\models t$ and $T_{R_{MR}} \not\models q$. □

Proposition 2.3.19. *For every profile $P \in \Phi^n$:*

- $T_{R_{MR}}(P) \subset T_{R_{MCSA}}(P)$,
- $T_{R_{MR}}(P) \subset T_{R_{MWA}}(P)$, *and*
- $T_{R_{MR}}(P) \subset T_{R_Y}(P)$.

Proof.

- $T_{R_{MR}}(P) \subset T_{R_{MCSA}}(P)$ is a consequence of $T_{R_{MR}}(P) \subset T_{R_{MSA}}(P)$, Proposition 2.3.15, and $T_{R_{MSA}}(P) \subset T_{R_{MCSA}}(P)$, Proposition 2.3.1.

- $T_{R_{MR}}(P) \subset T_{R_{MWA}}(P)$ is a consequence of $T_{R_{MR}}(P) \subset T_{R_{MSA}}(P)$, Proposition 2.3.15, and $T_{R_{MSA}}(P) \subseteq T_{R_{MWA}}(P)$, Proposition 2.3.2.

- $T_{R_{MR}}(P) \subset T_{R_Y}(P)$ is a consequence of $T_{R_{MR}}(P) \subset T_{R_{IY}}(P)$, Proposition 2.3.18, and $T_{R_{IY}}(P) \subset T_{R_Y}(P)$, Proposition 2.3.15.

To show that $T_{R_{MCSA}}(P) \not\subset T_{R_{MR}}(P)$, $T_{R_{MWA}}(P) \not\subset T_{R_{MR}}(P)$ and $T_{R_Y}(P) \not\subset T_{R_{MR}}(P)$ consider the profile P from Example 2.1.1.

- As it can be observed from Example 2.2.2, hence $T_{R_{MCSA}}(P) \models t$ and we showed in the proof of Proposition 2.3.18 that $T_{R_{MR}} \not\models t$.

- As it can be observed from Example 2.2.6, $R_{MWA}(P) = \{\{p \wedge r, p \wedge s, q, p \wedge q, t\}\}$, hence $T_{R_{MWA}}(P) \models t$ and $T_{R_{MR}} \not\models t$.

- As it can be observed from Example 2.2.6, $R_Y(P) = \{\{\neg(p \wedge r), \neg(p \wedge s), q, \neg(p \wedge q)\}\}$ hence $T_{R_{MCSA}}(P) \models q$ and we showed in the proof of Proposition 2.3.18 that $T_{R_{MR}} \not\models q$.

□

Proposition 2.3.20. *For every profile $P \in \Phi^n$, $T_{R_{MR}}(P) \subset T_{R_{RA}}(P)$.*

Proof. This inclusion is a consequence of $T_{R_{MR}}(P) \subset T_{R_{MSA}}(P)$, Proposition 2.3.18 and $T_{R_{MSA}}(P) \subset T_{R_{RA}}(P)$, Proposition 2.3.3. □

Proposition 2.3.21. *R_{MNAC} is incomparable with R_{MCSA}.*

Proof. To show that there exists a profile P such that $T_{R_{MCSA}}(P) \not\subset T_{R_{MNAC}}$, consider the pre-agenda and profile in the proof of Proposition 2.3.5 in which we have that $T_{R_{MCSA}}(P) \models \neg a$. There are 23 profiles Q at a minimal distance $D(P,Q) = 2$. We obtain $T_{R_{MNAC}}(P) \not\models \neg a$ because

$$\begin{aligned} R_{MNAC}(P) = \{ & \{a, a \to (b \vee c), \neg b, c, a \to (d \vee e), \neg d, e\}, \\ & \{a, a \to (b \vee c), \neg b, c, a \to (d \vee e), d, \neg e\}, \\ & \{a, a \to (b \vee c), b, \neg c, a \to (d \vee e), \neg d, e\}, \\ & \{a, a \to (b \vee c), b, \neg c, a \to (d \vee e), d, \neg e\}, \\ & \{a, \neg(a \to (b \vee c)), \neg b, \neg c, \neg(a \to (d \vee e)), \neg d, \neg e\} \\ & \{\neg a, a \to (b \vee c), \neg b, \neg c, a \to (d \vee e), \neg d, \neg e\}\}. \end{aligned}$$

To show that that there exists a profile P such that $T_{R_{MNAC}} \not\subset T_{R_{MCSA}}(P)$, consider the profile P from Example 2.1.1. We have $T_{R_{MNAC}}(P) \models q$, see Example 2.2.14, but $T_{R_{MCSA}}(P) \not\models q$, see Example 2.2.2. □

Proposition 2.3.22. *R_{MNAC} is incomparable with R_{MSA}.*

Proof. To show that there exists a P such that $T_{R_{MSA}}(P) \nsubseteq T_{R_{MNAC}}(P)$, consider the pre-agenda $\mathcal{A} = \{p, q, p \wedge q, p \wedge \neg q, \alpha_1, \alpha_2, q \wedge \neg p, \alpha_3, \alpha_4\}$, where
$\alpha_1 = p \wedge \neg q \wedge \neg q$,
$\alpha_2 = p \wedge \neg q \wedge \neg q \wedge \neg q$,
$\alpha_3 = q \wedge \neg p \wedge \neg p$,
$\alpha_4 = q \wedge \neg p \wedge \neg p \wedge \neg p$.
A profile for this pre-agenda is given in Table 2.15.

Voters	p	q	$p \wedge q$	$p \wedge \neg q$	α_1	α_2	$q \wedge \neg p$	α_3	α_4
$1 \times$	+	+	+	-	-	-	-	-	-
$1 \times$	+	-	-	+	+	+	-	-	-
$1 \times$	-	+	-	-	-	-	+	+	+
$M(P)$	+	+	-	-	-	-	-	-	-

Table 2.15: The profile P, counter-example for $T_{R_{MSA}}(P) \subset T_{R_{MNAC}}(P)$.

We obtain

$$R_{MSA}(P) = \{\{q, \neg(p \wedge q), \neg(p \wedge \neg q), \neg\alpha_1, \neg\alpha_2, \neg(q \wedge \neg p), \neg\alpha_3, \neg\alpha_4\}, \{p, \neg(p \wedge q), \neg(p \wedge \neg q), \neg\alpha_1, \neg\alpha_2, \neg(q \wedge \neg p), \neg\alpha_3, \neg\alpha_4\}, \{\neg(p \wedge q), \neg(p \wedge \neg q), \neg\alpha_1, \neg\alpha_2, \neg(q \wedge \neg p), \neg\alpha_3, \neg\alpha_4\}\}$$

Consequently $T_{R_{MSA}}(P) \models p \vee q$.

Voters	p	q	$p \wedge q$	$p \wedge \neg q$	α_1	α_2	$q \wedge \neg p$	α_3	α_4
$1 \times$	-	-	-	-	-	-	-	-	-
$1 \times$	+	-	-	+	+	+	-	-	-
$1 \times$	-	+	-	-	-	-	+	+	+
$M(P)$	-	-	-	-	-	-	-	-	-

Table 2.16: After changing the first three judgments of the first agent.

To obtain $R_{MNAC}(P)$, we need to change the first three judgments of the first voter, obtaining the profile given in Table 2.16. This is the minimal change, since if either the second or the third agent change either their judgment on p or their judgment on q, they have to change additional other three judgments. We obtain $R_{MNAC}(P) = \{\neg p, \neg q, \neg(p \wedge q), \neg(p \wedge \neg q), \neg\alpha_1, \neg\alpha_2, \neg(q \wedge \neg p), \neg\alpha_3, \neg\alpha_4\}$. We observe that $T_{R_{MNAC}}(P) \not\models p \vee q$.

To show that there exists a P such that $T_{R_{MNAC}}(P) \nsubseteq T_{R_{MSA}}(P)$, consider the profile P from Example 2.1.1. We have $T_{R_{MNAC}}(P) \models q$, see Example 2.2.14, but $T_{R_{MSA}}(P) \not\models q$, see Example 2.2.1. □

Proposition 2.3.23. *R_{MNAC} is incomparable with R_Y, R_{IY}, R_{RA}, R_{RY} and R_{MWA}.*

Proof. Consider the pre-agenda $\mathcal{A} = \{p, q, p \wedge q\}$ and the profile P from the proof of Proposition 2.3.11, given on Table 2.9. Since $R_{MNAC}(P) = M(P') \cup M(P'')$, where P' and P'' are as in Tables 2.17 and 2.18, we obtain that $\neg(p \wedge q) \in T_{R_{MNAC}}(P)$.

On the other hand, we obtain:

- $R_Y = \{\{p\}, \{q\}, \{\neg(p \wedge q)\}\}$,

Voters	p	q	$p \wedge q$
1×	+	+	+
1×	-	-	-
1×	-	+	-
$M(P)$	-	+	-

Table 2.17: The first profile, P', used to prove R_{MNAC} is incomparable with R_Y, R_{IY}, R_{RA}, R_{RY} and R_{MWA}.

Voters	p	q	$p \wedge q$
1×	+	+	+
1×	+	-	-
1×	-	-	-
$M(P)$	+	-	-

Table 2.18: The second profile, P'', used to prove R_{MNAC} is incomparable with R_Y, R_{IY}, R_{RA}, R_{RY} and R_{MWA}..

- $R_{IY} = \{\{p\},\{q\},\{\neg(p \wedge q)\}\}$
- $R_{RY} = \{\{p,q\},\{\neg q,\neg(p \wedge q)\},\{\neg p,\neg(p \wedge q)\}\}$
- $R_{RA} = \{\{p,\neg q,\neg(p \wedge q)\},\{p,\neg q,\neg(p \wedge q)\},\{p,q,p \wedge q\}\}$,
- $R_{MWA} = \{\{p,\neg q,\neg(p \wedge q)\},\{p,\neg q,\neg(p \wedge q)\},\{p,q,p \wedge q\}\}$.

Consequently $T_{R_{MNAC}}(P) \not\subset T_{R_Z}(P)$ for $Z \in \{Y, IY, RY, RA, MWA\}$.

To show that $T_{R_Y}(P) \not\subset T_{R_{MNAC}}(P)$ consider the profile P from Example 2.1.1. As it can be observed from Example 2.2.14, $T_{R_{MNAC}}(P) \models p \wedge r$, but we can observe in Example 2.2.10 that for this profile $T_{R_Y}(P) \models \neg(p \wedge r)$. Furthermore, we can observe in Example 2.2.11 that $T_{R_{IY}}(P) \not\models p \wedge r$; in Example 2.2.12 that $T_{R_{RY}}(P) \models \neg(p \wedge r)$ and in Example 2.2.8 that $T_{R_{RA}}(P) \models \neg(p \wedge r)$.

To show that $T_{R_{MWA}}(P) \not\subset T_{R_{MNAC}}(P)$, consider again the pre-agenda of the proof of Proposition 2.3.22 and its corresponding profile P given on Table 2.15. For this profile we get that $R_{MWA}(P) = \{p, q, p \wedge q, \neg(p \wedge \neg q), \neg\alpha_1, \neg\alpha_2, \neg(q \wedge \neg p), \neg\alpha_3, \neg\alpha_4\}$, since for this judgment set the weight is 17, and for the remaining three other possible judgment sets the weights are: 14 for the set of the judgment sets of the second, and third agent and 16 for the judgment set $\{\neg p, \neg q, \neg(p \wedge q), \neg(p \wedge \neg q), \neg\alpha_1, \neg\alpha_2, \neg(q \wedge \neg p), \neg\alpha_3, \neg\alpha_4\}$. Consequently, $T_{R_{MWA}} \models p \vee q$. In the proof of Proposition 2.3.22 we show that $T_{R_{MNAC}}(P) \not\models p \vee q$ for this profile. □

Proposition 2.3.24. *R_{MNAC} is incomparable with R_{MR}.*

Proof. To show $T_{R_{MR}}(P) \nsubseteq T_{R_{MNAC}}(P)$ consider the first part of the proof of Proposition 2.3.22. To show that $T_{R_{MNAC}}(P) \nsubseteq T_{R_{MR}}(P)$ consider the profile given in Table 2.13. We obtain $T_{R_{MNAC}}(P) \models \neg(p \wedge q)$ since $R_{MNAC}(P) = \{\{p, \neg q, \neg(p \wedge q)\}, \{\neg p, q \neg(p \wedge q)\}$. However $T_{R_{MR}}(P) \not\models \neg(p \wedge q)$, see the second part of the proof of Proposition 2.3.13. □

2.4 Conclusion

In this chapter we design judgment aggregation rules based on minimization. For a consensual group, a collective decision has to be such that it coincides with the view of the majority of the agents in the group. A consistent issue-majority set, in which each judgment is supported by a strict majority of agents, does not exist for every profile. The profiles for which such a set exists we call majority-consistent. We design judgment aggregation rules that minimally change the profile into a majority-consistent profile. When the profile is majority-consistent, no change is necessary. As a consequence the issue-majority set is always selected by our rules, when it exists. Each concept of minimal change gives rise to a new judgment aggregation rule.

We define ten judgment aggregation rules based on minimization, grouped in four families. We analyze how these rules relate to similar voting rules, but also to the judgment aggregation rules proposed in (Miller and Osherson, 2009). The aim of this chapter is to generate a large selection of concrete judgment aggregation rules that are majority-preserving and can be applied to any profile. Judgment aggregation theory normally follows the reverse methodology, studying the minimal sets of properties that can be simultaneously satisfied by a non-dictatorial or non-oligarchic rule or, such as the resent work of (Nehring et al., 2011; Nehring and Pivato, 2011), the characterization of rules which select from a desirable collection of judgment sets.

To determine if two judgment aggregation rules are distinct, we study the inclusion relations between the collective judgments selected by pairs of rules for the same profile. One purpose of the inclusion analysis, summarized in Table 2.4, is to verify whether two rules select different collective judgments for the same profile. Another purpose of this analysis is to qualify the rules to be able to distinguish them. Our analyses shows that the rules R_{MSA} and R_{MR} are very "weak" in the sense that they often select a very large number of judgment sets. In this sense the rule R_{IY} is weaker than R_Y.

The inclusion analysis enables us distinguish between the judgment aggregation rules based on the number of judgment sets they select. A consensual group usually needs only one collective judgment set to be selected by the judgment aggregation rule. Therefore the rules R_{MSA}, R_{IY} and R_{MR} are a bad choice for aggregation rules in consensual contexts. However we still need to be able to distinguish between the remaining rules and pair them with particular problems of decision reaching in consensual groups. To this end we return to these rules in Chapter 4 where we develop other properties for judgment aggregation rules and study how they are satisfied by R_{MCSA}, R_{RA}, R_{MWA}, R_{MNAC}, R_Y, R_{RY} and $R^{d_H,max}$.

Since the aim of application for our rules are computational contexts, one can also distinguish between rules by considering the complexity-theoretic properties of the rules. While we can reasonably expect that for some rules such as R_{RA}, finding the collective judgment sets can be done in a computationally efficient way, for other others such as the young rules we can expect that this task is not a problem of low computational complexity.

That the group decision minimizes the loss of information from the profile is only one way to interpret adherence to majority. What we considered in this chapter is the utilitarian perspective of minimizing loss of information. Another way is to minimize the loss of information from each individual judgment set in the profile, namely to take an egalitarian perspective. The rule $R^{d_H,max}$ in particular embodies this concept. An interesting class of rules can be constructed that minimally change each judgment set in the profile to obtain a majority-consistent profile. These types of rules would be of interests to groups of self-interested agents that need to reach a consensus on how to share a resource, the so called *fair division problems*, see for instance (Brams and Taylor, 1996, Introdcution).

3

Developing weighted ternary distance-based judgment aggregation rules

Abstract. Unlike in consensual groups, in hierarchical groups the adherence of the group decision to some majority is not the most relevant concern. The agent responsible for the decision in a hierarchical group needs to use the expertise of each agent that contributes opinions. Judgment weights can be used to represent the expertise of an agent regarding a given issue. While in consensual groups each agent can be expected to give a judgment on each issue, in hierarchical groups this is not necessarily the case. The aim of this chapter is to develop judgment aggregation rules for hierarchical groups. We extend the distance-based rules of the previous chapter into a class of judgment aggregation rules that aggregate three-valued judgments with associated weights. We give specific examples of rules and show the inclusion relationships between each pair. For this class of rules we also consider the computational complexity of the winner determination problem.

3.1 Introduction

Consider as an example of a hierarchical group a tourist recommender agent that needs to find the best hotel for you, provided your demands and conditions. This agent assembles information from various sources. While the hotel web page might be highly reliable on the issue of Wi-Fi being available in the rooms, the web page on user experience is the one with higher reliability than the hotel page when it comes to the issue of how silent the room is at night. There would be certain information that the tourist agent would disregard, for instance the quality of the bacon served for breakfast assessment from the vegetarian tourist blog. Also the agent is not going to be able to find information on all demands and conditions from every source. Consequently, the tourist recommender agent needs to use a judgment aggregation rule that aggregates judgment sets in which some agents abstain on some issues, *i.e.,* allow for three-valued judgments, and have different weights regarding the issues.

In many aggregation contexts for hierarchical groups, it is not feasible or desirable to request all the agents to vote on all the issues. In these contexts, different agents may have different levels of expertise on different issues and consequently their judgments should have a higher bearing on the collectively binding decisions. The rules introduced in Chapter 2 are defined for binary and unweight judgment sets. The problem pursued in this chapter is the development of weighted three-valued judgment aggregation rules.

Aggregation frameworks that allow for three-valued and even multi-valued judgments have been considered in for instance (Gärdenfors, 2006; Pauly and van Hees, 2006; Dietrich, 2007; Dokow and Holzman, 2010b; Li, 2010). Of these, only (Li, 2010) presents an actual rule, the sequential rule, for aggregating such judgments. Of all the rules introduced in Chapter 2, the rules based on the weighted majority graph and the rules based on distances can be extended to handle weights on judgments. From the rules based on the weighted majoritarian graph, we defined the rule R_{RA} and R_{MWA}. A weighted three-valued extension of R_{RA} can be easily constructed following the definitions and analysis of (Li, 2010). The rule R_{MWA}, as we showed, is equivalent to the distance-based rule $R_{d_H,\Sigma}$. Therefore, it is the class of distance-based rules the one we extend in this chapter. More precisely, we generalized further the family of $R_{d,\odot}$. The generalization approach we take can be directly applied to generalize R_{MNAC} into a weighted three-valued rule. Observe that if we wanted to consider only weights associated with agenda issues, we would be extending the rules based on the majority graph. If we wanted to consider only weights associated with an agent, the rules R_Y, R_{RY} and R_{IY} are the best candidates for extending.

The challenge in distance-based aggregation is not in aggregating multi-valued rules, but rather in aggregating rules in which weights are assigned to the judgments. A judgment is specified by a pair (agent, issue). Distance-based aggregation rules originate from belief merging (Konieczny and Pino-Pérez, 1999, 2002; Konieczny et al., 2004). Given a set of belief sets and a set of constraints, belief merging theory studies how to merge a set of belief bases in such a way that the resulting belief set, or sets, incorporate as much as possible from the individual beliefs and satisfy all the given constraints. In belief merging, considering weights for an agent is not uncommon; see for instance (Revesz, 1995). Weights assigned to an agent are also recently considered in judgment aggregation, (Nehring et al., 2011; Nehring and Pivato, 2011), however in neither field do we encounter weights assigned to an (agent,issue) pair, and weights assigned to issues are not considered in judgment aggregation. We solve the challenge of assigning weights to judgments by observing that sometimes a distance measure can itself be expressed using an arithmetic aggregator.

As in Chapter 2, here also we study inclusion relations between pairs of specific rules to verify that these rules select different judgments for one profile. The example scenario of a hierarchical group we consider in Chapter 5 is an example of agents making group decisions in uncertain environments. Since these agents are severely resource bounded we make a complexity-theoretic analysis for the family of aggregation rules we develop in this chapter.

This chapter is structured as follows. In Section 3.2 we introduce the necessary definitions. In Section 3.3 we design the family of weighted distance-based rules for aggregating ternary judgments and also give examples of specific rules in this family. In Section 3.4 we define an inclusion relation between judgment aggregation rules and analyze this relations between pairs of the specific rules introduced. Although we want a rule that aggregates ternary judgments, having only binary collective judgment sets selected by the rule can be desirable. In Section 3.5 we show how the family of weighted rules can be further modified to allow the decision-making agent to control structural properties of the selected (collective) judgment sets. The binary value of the collective judgment sets is such structural property. In Section 3.5 we also show how known judgment aggregation rules can be defined and extended when represented as a weighted distance-based rule. In Section 3.6 we show the computational complexity of determining whether a judgment set is among the ones selected from a distance-based merging rule. In Section 3.7 we make our conclusions.

3.2 Preliminaries

In this section we prepare the ground for building our family of extended distance based judgment aggregation rules. We construct a general judgment aggregation framework for representing three valued judgments. We also present the definitions of the family of rules we start from.

3.2.1 A dual framework for representing judgment aggregation problems

The problem of aggregating judgments was formulated by (List and Pettit, 2002) using logic representations. This problem, under the names of *abstract* or *algebraic aggregation* has precursors in (Gilbaud, 1966; Wilson, 1975) and (Rubinstein and Fishburn, 1986).

To represent an aggregation problem in a logic-based framework, one needs to specify a non-empty set $\mathcal{L}$ of well founded logic formulas and a binary (consequence) relation $\models\ \subseteq \mathcal{P}(\mathcal{L}) \times \mathcal{L}$, where $\mathcal{P}(\mathcal{L})$ denotes the power set of $\mathcal{L}$. $\mathcal{L}$ is called a *language* and its elements *propositions*. Propositions are not necessarily atomic formulas.

Definition 26. *A set of formulas $S \in \mathcal{L}$ is* logically interrelated *if there exists at least one $\varphi \in S$ such that either $S \backslash \{\varphi\} \models \varphi$ or $S \backslash \{\varphi\} \models \neg\varphi$.*

Definition 27. *A judgment aggregation problem is specified by a set of issues called an agenda $\mathcal{A} \subseteq \mathcal{L}$. Issues are the propositions on which the judgments are cast. The issues are interdependent, meaning that they share sub-formulas and/or are subject to an additionally specified set of constraints, $\mathcal{R} \subseteq \mathcal{L}$. The set $\mathcal{A} \cup \mathcal{R}$ is logically interrelated.*

A (binary) *judgment* on issue $a \in \mathcal{A}$ is usually defined, see for instance (Dietrich, 2007), as the choice of one element from the set $\{a, \neg a\}$. Pauly and van Hees (2006) construct a multi-valued logic framework in which a judgment is a valuation $v : \mathcal{A} \mapsto T$, where T is a set of values associated with gradient degrees of truth.

In an abstract framework no agenda is given, instead, the agents choose from a set of allowed binary sequences. For example, if the agenda of the aggregation problem in propositional logic were $\langle p, p \to q, q \rangle$, then the corresponding set of allowed sequences in an abstract framework would be $\{\langle 0,1,0 \rangle, \langle 0,1,1 \rangle, \langle 1,0,0 \rangle, \langle 1,1,1 \rangle\}$.

A dual framework for judgment aggregation with abstentions can be constructed: the judgments are represented both as propositions and as valuations. To this end, a ternary logic language $\mathcal{L}_3$ is used. We do not discuss here the possible logics $\mathcal{L}_3$ that can be used to represent the judgment aggregation problems and we do not concern ourselves with particular ternary logics. The choice of logic depends on the particular decision-problem that is modeled.

$\mathcal{L}_3$ is the set of well formed formulas of propositional logic $\mathcal{L}_{Prop}$ (in BNF):

$$\varphi ::= \top \mid \bot \mid p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \varphi \to \varphi \mid \varphi \leftrightarrow \varphi,$$

where $p \in \mathcal{L}_0$, $\mathcal{L}_0$ being the set of atoms.

The formulas of $\mathcal{L}_3$ are assigned values from the set $T = \{0, \frac{1}{2}, 1\}$. A valuation is a function $v_0 : \mathcal{L}_0 \mapsto T$, where the truth-values 0 and 1 are interpreted as in classical logic, $v_0(\bot) = 0$ and $v_0(\top) = 1$. The intermediate truth-value $\frac{1}{2}$ is interpreted depending on the semantics of the particular ternary logic $\mathcal{L}_3$ used. The semantics of $\mathcal{L}_3$ also determines how the function v_0 is extended to a function $v : \mathcal{L}_3 \mapsto T$.

A *judgment sequence* for an agenda $\mathcal{A}$, with cardinality m, is the sequence $A \in \{0, \frac{1}{2}, 1\}^m$ of judgments assigned to each of the issues in $\mathcal{A}$. We write $A(a)$ to denote the judgment assigned to $a \in \mathcal{A}$ according to sequence A. A *judgment set* for an agenda $\mathcal{A}$ is the set $\hat{A} \in 2^{\overline{\mathcal{A}}}$, where $\overline{\mathcal{A}} = \mathcal{A} \cup \{\neg a \mid a \in \mathcal{A}\}$. A judgment sequence A corresponds to a judgment set $\hat{A}$ and *vice versa*, if and only if, for all $a \in \mathcal{A}$ the value of a according to A is:

- 0 if and only if $\neg a \in \hat{A}$,
- 1 if and only if $a \in \hat{A}$, and
- $\frac{1}{2}$ if and only if $a \notin \hat{A}$ and $\neg a \notin \hat{A}$.

A consequence relation for a ternary logic, $\models_3$, is defined in the standard way (Urquhart, 2001). Given a set of formulas $\Gamma \subset \mathcal{L}_3$ and a formula $\psi \in \mathcal{L}_3$, we say that ψ is entailed by Γ, if and only if all assignments v that make Γ true, also make φ true. A formula ψ for which $\varnothing \models_3 \psi$ is a tautology of $\mathcal{L}_3$. A formula ψ is *satisfiable* in $\mathcal{L}_3$, if and only if there exists at least one valuation v such that $v(\psi) = 1$. A set of formulas Γ is *inconsistent* in $\mathcal{L}_3$ if and only if $\Gamma \models \bot$, and *consistent*, denoted $\Gamma \not\models \bot$ otherwise. Observe that Γ is consistent if there exists a valuation v such that $v(\bigwedge \Gamma) = 1$ or $v(\bigwedge \Gamma) = \frac{1}{2}$.

A judgment set $\hat{A}$ is *complete* when there exists no $a \in \mathcal{A}$ for which $a \notin A$ and $\neg a \notin A$. Correspondingly, a judgment sequence A is complete when there exists no $a \in \mathcal{A}$ for which $A(a) = \frac{1}{2}$.

Example 3.2.1 (Judgment sets and sequences)**.** *Consider an agenda $\mathcal{A} = \langle c_1, c_1 \to s_1, s_1 \rangle$, $\mathcal{R} = \varnothing$ and agents $N = \{1,2,3,4,5\}$. Let the judgment set for this agenda assigned by* 1 *and* 2 *be $\hat{A}_{1,2} = \{c_1, c_1 \to s_1, s_1\}$. The corresponding sequence for $\hat{A}_{1,2}$ is $A_{1,2} = \langle 1,1,1 \rangle$. Let the judgment set assigned by* 3, 4 *and* 5 *be $\hat{A}_{3,4,5} = \{\neg c_1\}$. The corresponding sequence for $\hat{A}_{3,4,5}$ is $A_{3,4,5} = \langle 0, \frac{1}{2}, \frac{1}{2} \rangle$. The judgment set and sequence for* 1 *and* 2 *are complete, while those for* 3, 4 *and* 5 *are not complete.*

A judgment set $\hat{A}$ and its corresponding sequence A are *consistent* for logic L_3, when $\hat{A} \cup \mathcal{R} \not\models \bot$. Given an agenda $\mathcal{A}$ and constraints $\mathcal{R}$, we can generate the set of all consistent judgment sets and corresponding sequences. The set of all consistent sequences A, with respect to $\mathcal{R}$, is $\mathbb{A}(\mathcal{A}, \mathcal{R}, \models_3)$, while the set of all corresponding consistent sets is $\hat{\mathbb{A}}(\mathcal{A}, \mathcal{R}, \models_3)$. To ease reading, we write simply $\mathbb{A}$ and $\hat{\mathbb{A}}$ whenever it is understandable from the context what $\mathcal{A}$, $\mathcal{R}$ and $\models_3$ are used. We denote by $\mathbb{A}^{\downarrow Prop}$ and by $\hat{\mathbb{A}}^{\downarrow Prop}$, the subsets of $\mathbb{A}$ and $\hat{\mathbb{A}}$ correspondingly, which satisfy the property $Prop$. For example, $Prop$ can be the subset of all judgment sequences from $\{0,1\}^m$; the subset of all judgment sequences in which the judgment on issue a is $\frac{1}{2}$ etc.

Example 3.2.2. *For agenda $\mathcal{A} = \{a_1, a_1 \wedge a_2, a_2\}$ and $\mathcal{R} = \varnothing$ the sets $\hat{\mathbb{A}}$ and $\mathbb{A}$ are:*

$$\begin{aligned}\mathbb{A} = \{ &\langle 0,0,0 \rangle, \langle \tfrac{1}{2},0,0 \rangle, \langle 0,0,\tfrac{1}{2} \rangle, \langle \tfrac{1}{2},0,\tfrac{1}{2} \rangle, \langle 1,0,0 \rangle, \langle 0,0,1 \rangle, \langle 1,0,\tfrac{1}{2} \rangle,\\ &\langle \tfrac{1}{2},0,1 \rangle, \langle 1,\tfrac{1}{2},1 \rangle, \langle 0,\tfrac{1}{2},0 \rangle, \langle \tfrac{1}{2},\tfrac{1}{2},0 \rangle, \langle 0,\tfrac{1}{2},\tfrac{1}{2} \rangle, \langle \tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2} \rangle, \langle 1,\tfrac{1}{2},0 \rangle,\\ &\langle 0,\tfrac{1}{2},1 \rangle, \langle 1,\tfrac{1}{2},\tfrac{1}{2} \rangle, \langle \tfrac{1}{2},\tfrac{1}{2},1 \rangle, \langle 1,1,1 \rangle \}\end{aligned}$$

$$\hat{\mathbb{A}} = \{\ \{\neg a_1, \neg(a_1 \wedge a_2), \neg a_2\}, \{\neg(a_1 \wedge a_2), \neg a_2\}, \{\neg a_1, \neg(a_1 \wedge a_2)\}, \{\neg(a_1 \wedge a_2)\},$$
$$\{a_1, \neg(a_1 \wedge a_2), \neg a_2\}, \{a_1, (a_1 \wedge a_2), \neg a_2\}, \{a_1, \neg(a_1 \wedge a_2)\}, \{\neg(a_1 \wedge a_2), \neg a_2\},$$
$$\{a_1, a_2\}, \{\neg a_1, \neg a_2\}, \{\neg a_2\}, \{\neg a_1\}, \{\}, \{a_1, \neg a_2\}, \{\neg a_1, a_2\}, \{a_1\}, \{a_2\},$$
$$\{a_1, (a_1 \wedge a_2), a_2\}\}$$

In judgment aggregation, a judgment profile is a structure that contains all the judgments made by agents N over the agenda items in $\mathcal{A}$. We give a dual definition of a profile: as a matrix of judgment and as a multiset of judgment sets. The profile defined as a matrix corresponds to the definition of a profile in abstract aggregation.

We define π to be a $n \times m$ matrix, where $n = |N|$ and $m = |\mathcal{A}|$. The elements of π are judgments: each row of the matrix is the judgment sets of one agent from N for all issues from $\mathcal{A}$, while each column contains the judgment sets of all of the agents from N for one issue from $\mathcal{A}$. We define an operator $\triangleright$ to retrieve a given row, and the operator $\triangledown$ to retrieve a given column from the matrix. Thus $\pi \triangleright i$ returns the sequence of all judgments made by agent i and $\pi \triangledown a$ returns a sequence of all values assigned to agenda issue a.

Definition 28 (Profile matrix)**.** *Let N be a set of n agents and $\mathcal{A}$ an agenda of m issues. A judgment profile $\pi \in \{0, \frac{1}{2}, 1\}^{n\times m}$ is a $|N| \times |\mathcal{A}|$-matrix $\pi = [p_{i,j}]$ where $p_{i,j} = v_i(a_j)$, and $i \in N$.*
The operators $\triangleright : \{0, \frac{1}{2}, 1\}^{n\times m} \times N \mapsto \{0, \frac{1}{2}, 1\}^m$ and $\triangledown : \{0, \frac{1}{2}, 1\}^{n\times m} \times \mathcal{A} \mapsto \{0, \frac{1}{2}, 1\}^n$ are defined as:

$\pi \triangleright i = \langle p_{i,j} \mid j \in \{1, \ldots, m\}\rangle$, *and*
$\pi \triangledown a_j = \langle p_{i,j} \mid i \in \{1, \ldots, n\}\rangle$.

Since the judgment sequence can be seen as a $1 \times m$ matrix, $A \triangledown a = A(a)$ denotes the value assigned to issue a according to the judgment sequence A. We use the notation $A_i = \pi \triangleright i$, and $p_{i,j}$ to denote the judgment $(\pi \triangleright i) \triangledown j$.

Example 3.2.3. *Consider the crew of cleaning robots $N = \{r_1, r_2, r_3\}$ that renders judgments on agenda $\mathcal{A} = \{p_1, p_2, p_3, g\}$ where:*

- p_1: *The meeting room is empty.*
- p_2: *The floors in the meeting room are dirty.*
- p_3: *There is garbage in the meeting room.*
- g: *The group should clean the meeting room.*

The constraint is that the group should clean the meeting room if and only if the room is empty and the floors are dirty or there is garbage in the room, i.e., $\mathcal{R} = \{(p_1 \wedge (p_2 \vee p_3)) \leftrightarrow g\}$. *One possible profile of judgments is:*

$$\pi = \begin{array}{c} \\ r_1 \\ r_2 \\ r_3 \end{array} \begin{array}{c} \begin{array}{cccc} p_1 & p_2 & p_3 & g \end{array} \\ \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & \frac{1}{2} & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

The judgment sequence of the robot r_2 is $\pi \triangleright r_2 = \langle 0, 1, \frac{1}{2}, 0\rangle$. The sequence of all judgments for p_3 is $\pi \triangledown p_3 = \langle 1, \frac{1}{2}, 0\rangle$.

Alternatively we define a profile to be a multi-set of judgment sets, multi-set since more than one agents can submit the same judgment set.

Definition 29 (Profile set). *Let N be a set of agents and $\mathcal{A}$ an agenda. A judgment profile P for $\mathcal{A}$ and N is a non-empty multiset of n judgment sets $A \in 2^{\overline{\mathcal{A}}}$.*

Example 3.2.4. *The profile P corresponding to profile π, from Example 3.2.3, is* $P = (\{p_1, p_2, p_3, g\}, \{\neg p_1, p_2, \neg g\}, \{p_1, \neg p_2, \neg p_3, \neg g\})$.

3.2.2 Binary unweight distance-based judgment aggregation rules

A judgment aggregation function is typically defined as $f(A_1, \ldots, A_n) \in \{0,1\}^n$, $A_1, \ldots, A_n \in \hat{\mathbb{A}}$, where $\hat{\mathbb{A}}$ is the set of all consistent and complete judgment sets. An abstract aggregation function is instead defined as $f : \{0,1\}^{m \times n} \mapsto \{0,1\}^m$. In the judgment aggregation literature it is always assumed, and we assume it here also, that $P \in \hat{\mathbb{A}}^n$ and that the allowed co-domain of f should also be $\hat{\mathbb{A}}$.

A judgment aggregation rule can be defined as $F : \mathbb{A}^n \mapsto \mathcal{P}(\mathbb{A})$, where $\mathcal{P}$ denotes the non-empty power set. The *distance-based procedure*, *DBP* defined in (Endriss et al., 2010b) is a judgment aggregation rule. We give the definition of this rule using our notation.

Let $\mathbb{A}^{\downarrow 01}$ denote the subset of $\mathbb{A}$ which includes only the sequences from $\{0,1\}^m$.

$$DBP(\pi) = \arg\min_{A \in \mathbb{A}^{\downarrow 01}} \sum_{i=1}^{n} \delta_H(A, \pi \triangleright i).$$

The *DBP* chooses the collective judgment sequences in the following way. First the *Hamming distances* δ_H between a judgment sequence $A \in \mathbb{A}^{\downarrow 01}$ and each of $\pi \triangleright i$ are calculated. A Hamming distance between two binary sequences is defined as

$$\delta_H(A, A') = \sum_{j=1}^{m} |A(a_j) - A'(a_j)|.$$

The rule selects those $A \in \mathbb{A}^{\downarrow 01}$ for which $\sum_{i=1}^{n} \delta_H(A, \pi \triangleright i)$ is minimal.

3.3 The judgment aggregation rules

The distance-based belief merging rules developed in (Konieczny and Pino-Pérez, 1999) are constructed by specifying a metric function (called a distance in the work in belief-merging) and an arithmetic aggregation function. In one direction, we generalize the *DBP* in the fashion of the operators of (Konieczny and Pino-Pérez, 1999), by considering a general aggregation function instead of $\sum_{i=1}^{n}$ and a general distance measure instead of δ_H.

If a judgment has assigned weight w we can see it as an unweighted judgment appearing w times in the profile, as if w agents gave the same judgment. Consequently, the aggregated judgments of *DBP* can be viewed as multiplied by a unique weight 1. We use this observation to generalize the *DBP* in another direction: before being aggregated the judgments are multiplied with their assigned weights.

3.3.1 Aggregation functions

Aggregation functions are defined in (Grabisch et al., 2009, pg.3). Since we use aggregation functions extensively, we give this definition here.

Definition 30. *Let $\mathbb{I}$ be a non-empty real interval. An aggregation function is a function*

$$\odot : \mathbb{I}^n \mapsto \mathbb{I}$$

that satisfies the following properties:

- *if $x \leqslant y$, then $\odot(x_1, \ldots, x, \ldots, x_n) \leqslant \odot(x_1, \ldots, y, \ldots, x_n)$ (non-decreasing);*
- *$\odot$ satisfies the boundary conditions:*
  - *$inf\ \odot = inf\ \mathbb{I}$;*
  - *$sup\ \odot = sup\ \mathbb{I}$.*

For example, the $\sum$ is an aggregation function defined for the interval $(-\infty, +\infty)$ since $\lim\limits_{x_i \to \pm\infty} \sum_{i=1}^{n} x_i = \pm\infty$.

We also include here the definitions of the most common properties of an aggregation function, as given in (Grabisch et al., 2009).

Definition 31. *An aggregation function is:*

- symmetric *if and only if $\odot(\mathbf{x}) = \odot([\mathbf{x}]_\sigma)$, for every $\mathbf{x} \in \mathbb{I}^n$ and permutation σ (Grabisch et al., 2009, pg.22);*
- associative *if and only if $\odot(x) = x$ for all $x \in \mathbb{I}$ and $\odot(\mathbf{x}, \odot(\mathbf{x}'), \mathbf{x}'') = \odot(\mathbf{x}, \mathbf{x}', \mathbf{x}'')$ for all $\mathbf{x}, \mathbf{x}', \mathbf{x}'' \in \bigcup_{n \in \mathbb{N}^0} \mathbb{I}^n$ (Grabisch et al., 2009, pg.22);*
- idempotent *if and only if $\odot(x, x \ldots, x) = x$ for all $x \in \mathbb{I}$ (Grabisch et al., 2009, pg.24).*

In (Konieczny and Pino-Pérez, 1999) the *minimality* of aggregation functions is considered. We give here the general definition of this property.

Definition 32. *An aggregation function $\odot$ satisfies minimality when $\odot(\mathbf{x}) = inf\ \mathbb{I}$ if and only if $\mathbf{x} = \odot(inf\ \mathbb{I}, \ldots, inf\ \mathbb{I})$.*

As a consequence of the infimum boundary condition on $\odot$ and the property of non-decreasing we have that if $\mathbf{x} = \odot(inf\ \mathbb{I}, \ldots, inf\ \mathbb{I})$ then $\odot(\mathbf{x}) = inf\ \mathbb{I}$. Therefore, $\odot$ satisfies minimality when if $\odot(\mathbf{x}) = inf\ \mathbb{I}$ then $\mathbf{x} = \odot(inf\ \mathbb{I}, \ldots, inf\ \mathbb{I})$.

We give the definitions of some common aggregation functions. The functions $\sum$, *Max*, and an operator *Gmax* are considered in (Konieczny and Pino-Pérez, 1999; Konieczny et al., 2004); $\sum$,Π, *max* and *AM* are presented in (Grabisch et al., 2009, pg.6). The arithmetic mean *AM* defined as $AM(x_1, \ldots, x_n) = \frac{1}{n} \sum(x_1, \ldots, x_n)$. Observe that, while it holds that $AM(\mathbf{x}) \geqslant AM(\mathbf{y})$ if and only if $\sum(\mathbf{x}) \geqslant \sum(\mathbf{y})$, the function *AM* is idempotent, while $\sum$ is not.

Definition 33. *For $\mathbf{x} \in \mathbb{I}^n$, the following functions are defined*

$$\begin{aligned}
\sum(\mathbf{x}) &= x_1 + \ldots + x_n;\\
max(\mathbf{x}) &= max(x_1, \ldots, x_n);\\
AM(\mathbf{x}) &= \tfrac{1}{n}\textstyle\sum_{i=1}^{n} x_i;\\
\Pi(\mathbf{x}) &= x_1 \cdot \ldots \cdot x_n;\\
Gmax(\mathbf{x}) &= \{(y_1, \ldots, y_n) \mid y_i \in \mathbf{x} \text{ and } y_1 \geqslant \cdots \geqslant y_n\}.
\end{aligned}$$

The functions Σ, *max*, *AM*, Π and *Gmax* are aggregation functions. The *Gmax* is also called a *leximax* operator. The $\sum$, *max*, *AM* and *Gmax* satisfy minimality on the interval $\mathbb{I} = \mathbb{R}^+ = [0, +\infty)$, while Π satisfies minimality on the interval $\mathbb{I} = [1, +\infty)$. To see that *Gmax* is an aggregation function, observe that *Gmax* sorts the input vector in a descending order. There is a one to one correspondence between the natural numbers and the sorted vector Konieczny et al. (2004).

All aggregation functions we present here are symmetric and satisfy associativity. Only the aggregation functions *AM* and *max* are idempotent.

3.3.2 Distance functions

Konieczny and Pino-Pérez (1999) define "distances" what Deza and Deza (2009) define to be a "metric". Here we follow the nomenclature and definitions of (Deza and Deza, 2009), primarily because we want to use a type of metric, not considered in (Konieczny and Pino-Pérez, 1999), that would enable us to construct weighted distance-based judgment aggregation rules. We present the definitions from (Deza and Deza, 2009, pg.3-4) and (Deza and Deza, 2009, pg.45) that we use.

Definition 34. *Let X be a set. A function $\delta : X \times X \to \mathbb{R}^+$ is called* a distance *on X if the following properties are satisfied for every $x, y, z \in X$:*

- $\delta(x,y) \geqslant 0$ *(non-negativity),*
- $\delta(x,y) = \delta(y,x)$ *(symmetry), and*
- $\delta(x,x) = 0$ *(reflexivity).*

A distance δ is called a metric *on X when for every $x, y, z \in X$:*

- $\delta(x,y) = 0$ *if and only if $x = y$ (identity of indiscernible);*
- $\delta(x,y) \leqslant \delta(x,z) + \delta(z,y)$ *(triangle inequality).*

The set (X, δ) is called a metric space when δ is a metric.

Definition 35. *Let $(X_1, d_1), (X_2, d_2), \ldots, (X_m, d_m)$ be a finite, or countable, number of metric spaces. A product metric d is a metric on the Cartesian product $X_1 \times X_2 \times \cdots \times X_m = \{x = (x_1, x_2, ..., x_m) : x_1 \in X_1, ..., x_m \in X_m\}$ defined as a function $\circledast$ of $\delta_1, ..., \delta_m$.*

Theorem 3.3.1. *If $X_1 = X_2 = \ldots X_n = X$, (δ, X) is a metric space and $\circledast$ is an aggregation function for $\mathbb{I} \in [0, +\infty)$ that satisfies minimality, then $d(\mathbf{x}, \mathbf{x}') = \circledast_{i=0}^{n} \delta(x_i, x_i')$ is a metric.*

Proof. Assume that δ is a metric. As a consequence $\delta(x_1,x_2) = 0$ if and only if $x_1 = x_2$, $\delta(x_1,x_2) = \delta(x_2,x_1)$ and $\delta(x_1,x_2) + \delta(x_2,x_3) \geqslant \delta(x_1,x_3)$ for any $x_1,x_2,x_3 \in X$. We need to show that d satisfies identity of indiscernible, symmetry and triangular inequality.

Identity of indiscernible
Since $\circledast$ satisfies minimality, then $d(\mathbf{x},\mathbf{x}') = \circledast_{i=0}^{n}\delta(x_i,x_i') = 0$ if and only if $\delta(x_i,x_i') = 0$ for each i. Therefore d satisfies the identity of indiscernible if δ satisfies this property.

Symmetry
From the definition $d(\mathbf{x},\mathbf{x}') = \circledast_{i=0}^{n}\delta(x_i,x_i')$, while $d(\mathbf{x}',\mathbf{x}) = \circledast_{i=0}^{n}\delta(x_i',x_i)$. We obtain that $d(\mathbf{x},\mathbf{x}') = d(\mathbf{x}',\mathbf{x})$ if $\delta(x_i,x_i') = \delta(x_i',x_i)$ for each i. Therefore d satisfies symmetry if δ satisfies symmetry.

Triangular inequality
Since δ satisfies triangular inequality, we have that $\delta(x_i^1,x_i^2) + \delta(x_i^2,x_i^3) \geqslant \delta(x_i^1,x_i^3)$ for each $i \in \{1,\ldots,m\}$. Consequently, $\circledast_{i=1}^{m}\delta(x_i^1,x_i^2) + \circledast_{i=1}^{m}\delta(x_i^2,x_i^3) \geqslant \circledast_{i=1}^{m}\delta(x_i^1,x_i^3)$ since both $\circledast$ and $+$ are non-decreasing. □

The well known functions, the Hamming distance and the *drastic distance*, are both product metrics that can be defined for any X. We will give their definitions, as well as introduce some other product metrics and distances. Some of these functions are defined only for $X = \{0,\frac{1}{2},1\}$, since we are interested in three-valued judgments.

Definition 36 (Hamming product metric)**.**

The Hamming metric is a function $\delta_H : X \times X \mapsto \{0,1\}$, *which indicates if two judgments differ. It is defined as:*

$$\delta_H(a_1,a_2) = \begin{cases} 0 & \textit{when } a_1 = a_2 \\ 1 & \textit{when } a_1 \neq a_2 \end{cases}.$$

The Hamming product metric d_H *is a function* $d_H : X^m \times X^m \mapsto \mathbb{N}^0$, *which indicates the number of judgments on which two sequences differ. It is defined as:*

$$d_H(\mathbf{x},\mathbf{x}') = \sum_{i=1}^{m}\delta_H(x_i,x_i').$$

Example 3.3.2. *Consider the agenda* $\mathcal{A} = \{a_1,a_2,a_3\}$ *and the sequences for it:* $A_1 = \langle 1,\frac{1}{2},0\rangle$, $A_2 = \langle \frac{1}{2},1,0\rangle$, $A_3 = \langle 1,\frac{1}{2},0\rangle$.

The Hamming metrics between these sequences are:

$d_H(A_1,A_2) = 2$ *because the judgments in* A_1 *and* A_2 *differ on issues* a_1 *and* a_2,

$d_H(A_2,A_3) = 2$ *also because the judgments in* A_2 *and* A_3 *differ on issues* a_1 *and* a_2,

$d_H(A_1,A_3) = 0$ *because the judgments in* A_1 *and* A_3 *are the same on all issues.*

The drastic distance between to sequences is one if the sequences are different and zero if they are the same.

Definition 37 (Drastic product metric)**.**

The drastic distance is a function $d_D : X^m \times X^m \mapsto \mathbb{N}^0$ *defined as:*

$$d_D(\mathbf{x},\mathbf{x}') = max(\delta_H(x_1,x'_1),\delta_H(x_2,x'_2),\ldots,\delta_H(x_m,x'_m)).$$

Example 3.3.3. *Consider the sequences* A_1*,* A_2 *and* A_3 *for the agenda from Example 3.3.2:* $A_1 = \langle 1,\frac{1}{2},0\rangle$*,* $A_2 = \langle 1,1,0\rangle$*,* $A_3 = \langle 1,\frac{1}{2},0\rangle$*.*

The drastic metrics between these sequences are:

$$d_D(A_1,A_2) = 1,$$

$$d_D(A_2,A_3) = 1,$$

$$d_D(A_1,A_3) = 0.$$

The Hamming distance does not make a difference by how much two judgments differ, but whether they differ or not. When $X = \{0,1\}$, this is not a problem, but for $X = \{0,\frac{1}{2},1\}$, we might want to consider by how much do two judgments differ. One way to capture this concept of distance is by the *Taxicab metric*, a measure introduced by Hermann Minkowski (1864-1909). The Taxicab metric between two judgment sequences is the sum of the absolute values of the difference between each judgment pairs in the sequences.

Definition 38 (Taxicab product metric)**.**

A taxicab metric is a function
$\delta_T : \{0,\frac{1}{2},1\} \times \{0,\frac{1}{2},1\} \mapsto \{0,\frac{1}{2},1\}$*, which indicates by how much do too judgments differ. It is defined as:*

$$\delta_T(x_1,x_2) = |x_1 - x_2|.$$

The Taxicab product metric is a function $d_T : \{0,\frac{1}{2},1\}^m \times \{0,\frac{1}{2},1\}^m \mapsto \mathbb{N}^0$ *defined as:*

$$d_T(\mathbf{x},\mathbf{x}') = \sum_{i=1}^{m} \delta_T(x_i,x'_i).$$

Observe that $\sum$ is an aggregation function that satisfies minimality on the interval $\mathbb{I} = \mathbb{N}^0$.

Example 3.3.4. *Consider* A_1*,* A_2 *and* A_3 *for the agenda from Example 3.3.2:* $A_1 = \langle 1,\frac{1}{2},0\rangle$*,* $A_2 = \langle 1,0,0\rangle$*,* $A_3 = \langle \frac{1}{2},1,0\rangle$*.*

The Taxicab metrics between these sequences are:

$$d_T(A_1,A_2) = |1-1| + |\tfrac{1}{2}-0| + |0-0| = \tfrac{1}{2},$$

$$d_T(A_2,A_3) = |1-\tfrac{1}{2}| + |0-1| + |0-0| = 1\tfrac{1}{2},$$

$$d_T(A_1,A_3) = |1-\tfrac{1}{2}| + |\tfrac{1}{2}-1| + |0-0| = 1.$$

Observation 3.3.5. *If* $A_1,A_2 \in \{0,1\}^m$ *then* $d_H(A_1,A_2) = d_T(A_1,A_2)$

The Taxicab metric does not make a difference whether the judgment is determined, 1/0 or an abstention. With choosing the metric the designer chooses how to treat the abstentions with respect to judgments “yes” and “no”. By choosing the Hamming or the drastic metric, the abstentions are treated as equal to the “yes” and “no” judgments. The “distance” functions can be defined to treat the abstentions differently.

The distance m_O assigns the distance zero from any judgment to an abstention, thus considering an abstention to equal to “yes” when compared to a “yes” judgment and “no” when compared to a “no” judgment.

Definition 39 (Optimistic metric)**.**

The optimistic distance is a function
$m_O : \{0, \frac{1}{2}, 1\}^m \times \{0, \frac{1}{2}, 1\}^m \mapsto \{0, 1\}$ *defined as*

$$m_O(\mathbf{x}, \mathbf{x}') = \sum_{i=1}^{m} \lfloor \delta_T(x_i, x_i') \rfloor.$$

Observation 3.3.6. *The function m_O satisfies non-negativity, symmetry and reflexivity, but it does not satisfy identity of indiscernible. The function $\circledast(\mathbf{x}) = (\sum \circ \lfloor\ \rfloor)(\mathbf{x})$ does not satisfy minimality.*

Example 3.3.7. *Consider the agenda from Example 3.3.2 and the judgment sets A_1, A_2 and A_3 for this agenda: $A_1 = \langle 1, \frac{1}{2}, 0 \rangle$, $A_2 = \langle \frac{1}{2}, \frac{1}{2}, 0 \rangle$, $A_3 = \langle 0, \frac{1}{2}, 0 \rangle$.*

The optimistic metrics between these sequences are:

$$m_O(A_1, A_2) = 0 + 0 + 0 = 0,$$

$$m_O(A_2, A_3) = 0 + 0 + 0 = 0,$$

$$m_O(A_1, A_3) = 1 + 0 + 0 = 1.$$

In all the metrics we presented, the number assigned to a pair of judgment sequences is always obtained by comparing only the two sequences in the pair. (Duddy and Piggins, 2011) introduce a more complex metric for complete judgment sets as a function $g : \mathbb{A}^{\downarrow 01} \times \mathbb{A}^{\downarrow 01}$, that is not a product metric. Their metric is defined in the following way. Let $\mathcal{G} = (\mathbb{A}^{\downarrow 01}, \mathcal{E})$ be a graph where the vertices are the judgment sequences in $\mathbb{A}^{\downarrow 01}$. The set of edges $\mathcal{E} \in \mathbb{A}^{\downarrow 01} \times \mathbb{A}^{\downarrow 01}$ consists of pairs $(A_1, A_2) \in \mathcal{E}$ for which there exists no $A \in \mathbb{A}^{\downarrow 01}$ such that $d_H(A_1, A_2) = d_H(A_1, A) + d_H(A, A_2)$. A metric $g(A_1, A_2)$ is the number of edges in the shortest path between A_1 and A_2.

We can extend the metric of (Duddy and Piggins, 2011) to incomplete judgments by using a graph $\mathcal{G}_3 = (\mathbb{A}, \mathcal{E})$ and allowing for an edge to exist between A_1 and A_2 if an only if there exists no $A \in \mathbb{A}$ such that $d_T(A_1, A_2) = d_T(A_1, A) + d_T(A, A_2)$. We call this metric d_G. We calculate $g(A_1, A_2)$ as the number of edges in the shortest path between A_1 and A_2. However, whether this metric is meaningful depends on the semantics of the ternary logics. For instance, for the logics of Kleeney or Łukasiewicz, and a classical $\models_3$, there is a judgment sequence at a Taxicab distance $\frac{1}{2}$ or at a Hamming distance 1 for each judgment sequence in $\mathbb{A}$. This is because a judgment set $\hat{A}$ is consistent when $\hat{A} \cup \mathcal{R} \not\models_3 \perp$ is false or unknown to be false, *i.e.*, evaluated to $\frac{1}{2}$.

Other distances and metrics can be defined.

3.3.3 Weights

Different agents may not be equally competent, or able, to give judgments on all agenda issues. Enriching the judgment aggregation problem representation with a collection of weights captures this variety. There are three possible types of weights that can be considered: weight associated with an agent, weight associated with an agenda issue and weight associated with a judgment, *i.e.*, with a $(agent, issue)$ pair. All types of weights can be represented with a *weight matrix*. Given a set of agents N and an agenda $\mathcal{A}$, a weight matrix W is a $[w_{i,j}]_{n\times m}$ matrix. The elements of W, $w(i,j) \in \mathbb{R}^+$, are the weights assigned to the judgments given by $i \in N$ for an $a_j \in A$.

One interpretation of the judgment weights is that of the weight representing reputation or perceived accuracy of an agent i regarding issue a at a given time t. The reputation can be defined simply as the ratio between the number of times an agent is asked to make a judgment on issue a_j before time moment t and the number of times, until t has his judgment been confirmed. Assume that $r(i,j,t) \in [0,1]$ is the normalized reputation of agent i regarding $a_j \in \mathcal{A}$. Weights can be constructed from reputation $r(i,j,t)$ as $w_{i,j}(t) = 1 + r(i,j,t)$, thus maintaining that $w_{i,j} \geqslant 1$. When the reputation of the agent is 0, namely none of his judgments is confirmed, his weight is 1, because the opinion of this agent still needs to be considered.

The cases when no weights are supplied, when weights associated with an agent are supplied, or when weights associated with an agenda issue are supplied, can all be represented as a special case of W. If no weights are given, then $W = U$, where U is such that for each i and j, $w_{i,j} = 1$.

If the weights associated with an agent are given then for each i, $w_{i,1} = w_{i,2} = \cdots = w_{i,m}$. In judgment aggregation problems that use this type of weights, the reputation of the agents is set beforehand and does not depend on the agenda. *E.g.*, for a set of three agents and an agenda of three issues the matrix W is a possible agent weight matrix.

$$W = \begin{bmatrix} 1 & 1 & 1 \\ 1.2 & 1.2 & 1.2 \\ 0.2 & 0.2 & 0.2 \end{bmatrix}$$

If the weights associated with agenda issues are given then for each j, $w_{1,j} = w_{2,j} = \cdots = w_{n,j}$. This type of weights distinguishes between the relevance of one issue over another. These weights do not depend on the agent who renders a judgment. If an issue a is more relevant then issue a', then the difference in judgments on a is more severe than the difference in judgments on a'. *E.g.*, for a set of three agents and an agenda of three issues the matrix W is a possible issue weight matrix.

$$W = \begin{bmatrix} 1 & 2 & 1.5 \\ 1 & 2 & 1.5 \\ 1 & 2 & 1.5 \end{bmatrix}$$

Consider the so called “truth-functional” agendas, which can be partitioned into a set of premises and a set of conclusions. Based on this partition, one can distinguish between premise-based rules, which place higher importance on the premises and conclusion-based rules that place higher importance on the conclusions. According to the premise-based aggregation rule defined in (Dietrich and Mongin, 2010), the collective judgments on the issues from the set of premises are the judgments supported by a strict majority. We can use the weights to force a rule aggregator to be premise-based, by increasing the weights on the premises, or conclusion based, by increasing the weights on the conclusions.

In this thesis we work under the assumption that the weights in W are specified by the agent who aggregates the judgments. The presence of judgment weights in an aggregation problem implies that one pre-established agent or service aggregates the judgments centrally. This implication is due to the collective judgments selected depending on who assigns the associated weights.

Lastly the weights can be used to represent aggregation problems in which not all agents are allowed to give judgments on all agenda issues. If the aggregating agent is not interested in the judgment on a_j of agent i, then he should set $w(i,j) = 0$. The zero weight can also be used in the case when the agents fail to report a judgment on a given issue due to for instance technical difficulties in communication.

3.3.4 Distance-based rules, the generalization

We can now "lift" the definition of the premise-based procedure along the two directions and construct a new family of weighted distance-based judgment aggregation rules.

Definition 40. *Let $\mathcal{A} = \{a_1, \ldots, a_m\}$ be an agenda, $\mathcal{R}$ a set of constraints, N a set of agent names, and $\mathbb{A}(\mathcal{A}, \mathcal{R}, \models_3)$ the set of all consistent three-valued judgment sequences for $\mathcal{A}$ and $\mathcal{R}$. Let $\odot$ be an aggregation function, and d a product metric. The metric d is constructed from an aggregation function $\circledast$ that satisfies minimality and a distance δ. A weighted distance-based aggregation rule is a function $\Delta^{d,\odot} : \mathbb{A}^n \times (\mathbb{R}^+)^{n\times m} \mapsto \mathcal{P}(\mathbb{A})$, defined as:*

$$\Delta^{d,\odot}(\pi, W) = \arg\min_{A \in \mathbb{A}} \ \odot_{i=1}^{n} \circledast_{j=1}^{m} w(i,j) \cdot \delta(A(a_j), \pi_{i,j}).$$

Example 3.3.8. *Consider the profile π from Example 3.4.1 for agents $N = \{1,2,3\}$. Let the weight matrix be W.*

$$\pi = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} a_1 a_2 a_3 \\ \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{array} \qquad W = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} a_1 a_2 a_3 \\ \begin{bmatrix} 1 & \frac{3}{2} & 2 \\ 1 & \frac{4}{3} & \frac{5}{4} \\ 1 & \frac{3}{2} & \frac{7}{4} \end{bmatrix} \end{array}$$

We use $\odot = \circledast = \sum$ and $\delta = \delta_T$. The sum of weighted distances between $A_1 = \pi \rhd 1$, and *each sequence of π, when $\circledast = \sum$ is:*

$$
\begin{aligned}
&\textstyle\sum_{i=1}^{3}\sum_{j=1}^{3} w(i,j)\cdot\delta_T(A_1\triangledown a_j, A_i\triangledown a_j)\\
&= \textstyle\sum_{i=1}^{3}(w(i,1)\cdot\delta_T(A_1\triangledown a_1, A_i\triangledown a_1)\\
&\qquad +w(i,2)\cdot\delta_T(A_1\triangledown a_2, A_i\triangledown a_2)\\
&\qquad +w(i,3)\cdot\delta_T(A_1\triangledown a_3, A_i\triangledown a_3))\\
&= (w(1,1)\cdot\delta_T(A_1\triangledown a_1,A_1\triangledown a_1)+w(1,2)\cdot\delta_T(A_1\triangledown a_2,A_1\triangledown a_2)+w(1,3)\cdot\delta_T(A_1\triangledown a_3,A_1\triangledown a_3))\\
&+ (w(2,1)\cdot\delta_T(A_1\triangledown a_1,A_2\triangledown a_1)+w(2,2)\cdot\delta_T(A_1\triangledown a_2,A_2\triangledown a_2)+w(2,3)\cdot\delta_T(A_1\triangledown a_3,A_2\triangledown a_3))\\
&+ (w(3,1)\cdot\delta_T(A_1\triangledown a_1,A_3\triangledown a_1)+w(3,2)\cdot\delta_T(A_1\triangledown a_2,A_3\triangledown a_2)+w(3,3)\cdot\delta_T(A_1\triangledown a_3,A_3\triangledown a_3))\\
&= (1\cdot|1-1|+\tfrac{3}{2}\cdot|\tfrac{1}{2}-\tfrac{1}{2}|+2\cdot|\tfrac{1}{2}-\tfrac{1}{2}|\\
&+ 1\cdot|1-1|+\tfrac{4}{3}\cdot|\tfrac{1}{2}-0|+\tfrac{5}{4}\cdot|\tfrac{1}{2}-1|\\
&+ 1\cdot|1-0|+\tfrac{3}{2}\cdot|\tfrac{1}{2}-0|+\tfrac{7}{4}\cdot|\tfrac{1}{2}-1|\\
&= 0+0+0+0+\tfrac{4}{6}+\tfrac{5}{8}+\tfrac{6}{8}+\tfrac{3}{4}+\tfrac{7}{8}\\
&= 3.66
\end{aligned}
$$

If the weight matrix contains weights associated with an agent, we can define a weighted aggregation rule without the requirement that d is a product of distances. Let V be a weight tuple $V = [w]_{n\times 1}$ containing the weight of each agent.

Definition 41. *An agent-weighted distance-based aggregation rule is a function $\Delta_V^{d,\odot} : \mathbb{A}^n \times (\mathbb{R}^+)^n \mapsto \mathcal{P}(\mathbb{A})$, defined as:*

$$\Delta_V^{\delta,\odot}(\pi, V) = \operatorname*{arg\,min}_{A\in\mathbb{A}} \ \odot(w_1\cdot d(A,\pi\triangleright 1),\ldots,w_n\cdot d(A,\pi\triangleright n)).$$

The co-domain of the rules Δ must be a power set of $\mathbb{A}$. However, we can define the rules Δ for a profile $\pi \in \{0,\frac{1}{2},1\}^{n\times m}$ instead of $\pi \in \mathbb{A}^n$, without much modification. This means that the distance-based judgment aggregation rules can be applied to sequences which are not consistent for the chosen logic. Miller (2008) studied the case when each agent is allowed to use his own *subjective rules* $\mathcal{R}_i$ for the judgments he produces. In addition to this variation, one can also conceive the case when each agent uses individual *subjective semantics*. Our rules $\Delta^{d,\odot}$ can be applied to both of these two cases. The definition of $\Delta^{d,\odot}$ does not explicitly consider the ternary logic semantics. This concern is resolved by defining the set $\mathbb{A}$. The difference between aggregating sequences consistent in for instance Post logic (Post, 1921) and Kleeney logic (Kleene, 1938) is that the co-domain of $\Delta^{d,\odot}$ is different for each of these logics.

3.4 (Non)Inclusion relationships between specific rules

Each combination of $\odot, \circledast$ and δ gives rise to another aggregation rule, however not all of these rules are meaningful. We call a rule *meaningless* if for every profile, except the profile in which $\pi_1 = \pi_2 = \cdots = \pi_n$, each judgment set that is in the profile is also selected as a collective judgment set. Namely, a rule is meaningless when for all $\pi \in \mathbb{A}^n$ and for all $i \in N$, $\pi\triangleright i \in \Delta^{d,\odot}(\pi, U)$. For instance, combining Π with any distance function gives rise to a meaningless rule, since for each $\pi \in \mathbb{A}^n$ we obtain that $\Delta^{d,\Pi}(\pi, U) = \pi$. Combining *max* with d_D gives rise to a meaningless rule as well, since unless $\pi_1 = \pi_2 = \cdots = \pi_n$ we obtain $\Delta^{d,\Pi}(\pi, U) = \mathbb{A}$.

Combining *max* with d_h or d_t we obtain a rule that behaves as a plurality voting rule, namely it selects the judgment sequence that is supported by the largest number of agents, regardless of how big this number is with respect to the total number of agents n. When such a sequence does not exist, the entire profile is selected.

To use Π we need to use a function whose domain is $[1,+\infty)$ instead of $[0,+\infty)$. One such function is $m_P : \{0,1\}^m \times \{0,1\}^m \mapsto \mathbb{N}$ defined as:

$$m_P(\mathbf{x},\mathbf{x}') = \Pi_{i=1}^{m} 2^{\delta_H(x_1,x_i')}.$$

We can then obtain an operator $\Delta^{m_P,\Pi}$. However, for every profile π, $\Delta^{m_P,\Pi}(\pi,W)$ selects the same judgment sets as $\Delta^{d_H,\Sigma}(\pi,W)$: it is enough to observe that for any three judgment sequences A, A_1 and A_2, from $\{0,\frac{1}{2},1\}^m$ it holds

$$\sum_{i=1}^{m} w_{1,i} \cdot \delta_H(A(i),A_1(i)) \leqslant \sum_{i=1}^{m} w_{1,i} \cdot \delta_H(A(i),A_2(i))$$

if and only if

$$\prod_{i=1}^{m} w_{1,i} \cdot 2^{\delta_H(A(i),A_1(i))} \leqslant \prod_{i=1}^{m} w_{1,i} \cdot 2^{\delta_H(A(i),A_2(i))}.$$

We can construct a judgment aggregation function for the interval $\mathbb{I} = [1,+\infty)$ as

$$\Pi^*(\mathbf{x}) = \Pi_{i=1}^{n}(x_i+1).$$

Using $\Pi^* = \Pi \circ g$, where $g(x) = x+1$, we can obtain meaningful judgment aggregation rules.

We can illustrate the specific rules that can be obtained with the metrics and aggregation functions that we introduced through an example.

Example 3.4.1. *Consider the agenda $\mathcal{A}$ and corresponding set $\mathbb{A}$ from Example 3.2.2, and the profile:*

$$\pi = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Table 3.1 gives the results for $\Delta^{d_H,\odot}(\pi,U)$, Table 3.2 gives the results for $\Delta^{d_D,\odot}(\pi,U)$, Table 3.3 gives the results for $\Delta^{d_T,\odot}(\pi,U)$, and Table 3.4 gives the results for $\Delta^{m_O,\odot}(\pi,U)$. The pink fields are the minima in each column corresponding to an aggregation rule. For this π, all rules select $\langle 1,0,0\rangle$. However this is not the case with all profiles.

A	$d_H(A,\langle 1,\frac{1}{2},\frac{1}{2}\rangle)$	$d_H(A,\langle 1,0,0\rangle)$	$d_H(A,\langle 0,0,0\rangle)$	$\sum$	Max	Gmax	Π*	AM
$\langle 0,0,0\rangle$	3	1	0	4	3	(3,1,0)	8	1.33
$\langle \frac{1}{2},0,0\rangle$	3	1	1	5	3	(3,1,1)	16	1.66
$\langle 0,0,\frac{1}{2}\rangle$	2	2	2	6	2	(2,2,2)	27	2
$\langle \frac{1}{2},0,\frac{1}{2}\rangle$	2	1	2	5	2	(2,2,1)	18	1.66
$\langle 1,0,0\rangle$	2	0	1	3	2	(2,1,0)	6	1
$\langle 0,0,1\rangle$	3	2	1	6	3	(3,2,1)	24	2
$\langle 1,0,\frac{1}{2}\rangle$	1	1	2	4	2	(2,1,1)	12	1.33
$\langle \frac{1}{2},0,1\rangle$	3	2	2	7	3	(3,2,2)	36	2.33
$\langle 1,\frac{1}{2},1\rangle$	1	1	3	7	3	(3,3,1)	16	2.33
$\langle 0,\frac{1}{2},0\rangle$	2	2	1	5	2	(2,2,1)	18	1.66
$\langle \frac{1}{2},\frac{1}{2},0\rangle$	2	2	2	6	2	(2,2,2)	27	2
$\langle 0,\frac{1}{2},\frac{1}{2}\rangle$	1	3	2	6	3	(3,2,1)	24	2
$\langle \frac{1}{2},\frac{1}{2},\frac{1}{2}\rangle$	1	3	3	7	3	(3,3,1)	32	2.33
$\langle 1,\frac{1}{2},0\rangle$	1	1	2	4	2	(2,1,1)	12	1.33
$\langle 0,\frac{1}{2},1\rangle$	2	3	2	7	3	(3,2,2)	36	2.33
$\langle 1,\frac{1}{2},\frac{1}{2}\rangle$	0	2	3	5	3	(3,2,0)	12	1.66
$\langle \frac{1}{2},\frac{1}{2},1\rangle$	2	3	3	8	3	(3,2,2)	48	2.66
$\langle 1,1,1\rangle$	2	2	3	7	3	(3,2,2)	36	2.33

Table 3.1: The Hamming metric between the sequences in π and the elements of $\mathbb{A}$.

A	$d_D(A,\langle 1,\frac{1}{2},\frac{1}{2}\rangle)$	$d_D(A,\langle 1,0,0\rangle)$	$d_D(A,\langle 0,0,0\rangle)$	$\sum$	Max	Gmax	Π^*	AM
$\langle 0,0,0\rangle$	1	1	0	2	1	(1,1,0)	4	0.66
$\langle \frac{1}{2},0,0\rangle$	1	1	1	3	1	(1,1,1)	8	1
$\langle 0,0,\frac{1}{2}\rangle$	1	1	1	3	1	(1,1,1)	8	1
$\langle \frac{1}{2},0,\frac{1}{2}\rangle$	1	1	1	3	1	(1,1,1)	8	1
$\langle 1,0,0\rangle$	1	0	1	2	1	(1,1,0)	4	0.66
$\langle 0,0,1\rangle$	1	1	1	3	1	(1,1,1)	8	1
$\langle 1,0,\frac{1}{2}\rangle$	1	1	1	3	1	(1,1,1)	8	1
$\langle \frac{1}{2},0,1\rangle$	1	1	1	3	1	(1,1,1)	8	1
$\langle 1,\frac{1}{2},1\rangle$	1	1	1	3	1	(1,1,1)	8	1
$\langle 0,\frac{1}{2},0\rangle$	1	1	1	3	1	(1,1,1)	8	1
$\langle \frac{1}{2},\frac{1}{2},0\rangle$	1	1	1	3	1	(1,1,1)	8	1
$\langle 0,\frac{1}{2},\frac{1}{2}\rangle$	1	1	1	3	1	(1,1,1)	8	1
$\langle \frac{1}{2},\frac{1}{2},\frac{1}{2}\rangle$	1	1	1	3	1	(1,1,1)	8	1
$\langle 1,\frac{1}{2},0\rangle$	1	1	1	3	1	(1,1,1)	8	1
$\langle 0,\frac{1}{2},1\rangle$	1	1	1	3	1	(1,1,1)	8	1
$\langle 1,\frac{1}{2},\frac{1}{2}\rangle$	0	1	1	2	1	(1,1,0)	4	0.66
$\langle \frac{1}{2},\frac{1}{2},1\rangle$	1	1	1	3	1	(1,1,1)	8	1
$\langle 1,1,1\rangle$	1	1	1	3	1	(1,1,1)	8	1

Table 3.2: The drastic metric between a sequence in π and the elements of $\mathbb{A}$.

If a rule is not meaningless, how can we determine if it selects the same collective judgment sequences as another rule? In Chapter 2 we defined the equivalence and set inclusion of judgment aggregation rules through the logical theory of the rules. The majority of the rules based on minimization select incomplete judgment sets. The logical theory based comparison is adequate there since it compares the collective judgments selected by the two rules that are being considered.

The distance-based judgment aggregation rules always select at least one collective judgment for each agenda issue, therefore a more adequate relation analysis between the rules is one that considers which judgment sequences as a whole are selected and not the individual judgments. We introduce a variant of rule relations, considering a rule to be more discriminant than another rule when the set of sequences selected by the first rule always includes the set of sequences selected by the second rule, under the same profile and weight matrix. We define this rule relations concept formally.

Definition 42 (Rule Relations)**.** *Let F_1 and F_2 be two judgment aggregation rules defined as $F_1 : S^n \times (\mathbb{R}^+)^{n\times m} \mapsto \mathcal{P}(S)$ and $F_2 : S^n \times (\mathbb{R}^+)^{n\times m} \mapsto \mathcal{P}(S)$.*

We say that rule F_1 is included in rule F_2, denoted $F_1 \subset F_2$, if for every $\pi \in S^n$ and $W \in (\mathbb{R}^+)^{n\times m}$ it holds that $F_1(\pi,W) \subset F_2(\pi,W)$.

A rule F_1 is incomparable with rule F_2, denoted $F_1 \not\approx F_2$, if there exists a pair $\pi \in S^n$ and $W \in (\mathbb{R}^+)^{n\times m}$ such that $F_1(\pi,W) \not\subset F_2(\pi,W)$ and $F_2(\pi,W) \not\subset F_1(\pi,W)$.

A rule F_1 is equal to rule F_2, denoted $F_1 = F_2$, if for every $\pi \in S^n$ and $W \in (\mathbb{R}^+)^{n\times m}$ it holds that $F_1(\pi,W) = F_2(\pi,W)$.

A	$d_T(A,\langle 1,\frac{1}{2},\frac{1}{2}\rangle)$	$d_T(A,\langle 1,0,0\rangle)$	$d_T(A,\langle 0,0,0\rangle)$	$\sum$	Max	Gmax	Π^*	AM
$\langle 0,0,0\rangle$	2	1	0	3	2	(2,1,0)	6	1
$\langle \frac{1}{2},0,0\rangle$	$1\frac{1}{2}$	$1\frac{1}{2}$	$\frac{1}{2}$	$3\frac{1}{2}$	$1\frac{1}{2}$	$(1\frac{1}{2},\frac{1}{2},\frac{1}{2})$	9.375	1.16
$\langle 0,0,\frac{1}{2}\rangle$	$1\frac{1}{2}$	$1\frac{1}{2}$	$\frac{1}{2}$	$3\frac{1}{2}$	$1\frac{1}{2}$	$(1\frac{1}{2},1\frac{1}{2},\frac{1}{2})$	9.375	1.16
$\langle \frac{1}{2},0,\frac{1}{2}\rangle$	1	1	1	3	1	(1,1,1)	8	1
$\langle 1,0,0\rangle$	1	0	1	2	1	(1,1,0)	4	0.6
$\langle 0,0,1\rangle$	2	2	1	5	2	(2,2,1)	18	1.6
$\langle 1,0,\frac{1}{2}\rangle$	$\frac{1}{2}$	$\frac{1}{2}$	$1\frac{1}{2}$	$2\frac{1}{2}$	$1\frac{1}{2}$	$(1\frac{1}{2},\frac{1}{2},\frac{1}{2})$	5.645	0.83
$\langle \frac{1}{2},0,1\rangle$	$1\frac{1}{2}$	$1\frac{1}{2}$	$1\frac{1}{2}$	$4\frac{1}{2}$	$1\frac{1}{2}$	$(1\frac{1}{2},1\frac{1}{2},1\frac{1}{2})$	15.625	1.5
$\langle 1,\frac{1}{2},1\rangle$	$\frac{1}{2}$	$1\frac{1}{2}$	$2\frac{1}{2}$	$4\frac{1}{2}$	$2\frac{1}{2}$	$(2\frac{1}{2},1\frac{1}{2},\frac{1}{2})$	13.125	1.5
$\langle 0,\frac{1}{2},0\rangle$	$1\frac{1}{2}$	$1\frac{1}{2}$	$\frac{1}{2}$	$3\frac{1}{2}$	$1\frac{1}{2}$	$(1\frac{1}{2},1\frac{1}{2},\frac{1}{2})$	9.375	1.16
$\langle \frac{1}{2},\frac{1}{2},0\rangle$	1	1	1	3	1	(1,1,1)	8	1
$\langle 0,\frac{1}{2},\frac{1}{2}\rangle$	1	2	1	4	2	(2,1,1)	12	1.3
$\langle \frac{1}{2},\frac{1}{2},\frac{1}{2}\rangle$	$\frac{1}{2}$	$1\frac{1}{2}$	$1\frac{1}{2}$	$3\frac{1}{2}$	$1\frac{1}{2}$	$(1\frac{1}{2},1\frac{1}{2},\frac{1}{2})$	9.375	1.16
$\langle 1,\frac{1}{2},0\rangle$	$\frac{1}{2}$	$\frac{1}{2}$	$1\frac{1}{2}$	$2\frac{1}{2}$	$1\frac{1}{2}$	$(1\frac{1}{2},\frac{1}{2},\frac{1}{2})$	5.625	0.83
$\langle 0,\frac{1}{2},1\rangle$	$1\frac{1}{2}$	$2\frac{1}{2}$	$1\frac{1}{2}$	$5\frac{1}{2}$	$2\frac{1}{2}$	$(2\frac{1}{2},1\frac{1}{2},1\frac{1}{2})$	21.875	1.83
$\langle 1,\frac{1}{2},\frac{1}{2}\rangle$	0	1	2	3	2	(2,1,0)	6	1
$\langle \frac{1}{2},\frac{1}{2},1\rangle$	1	2	2	5	2	(2,2,1)	18	1.66
$\langle 1,1,1\rangle$	1	2	3	6	3	(3,2,1)	24	2

Table 3.3: The taxicab metric between a sequences in π and the elements of $\mathbb{A}$.

A	$m_O(A,\langle 1,\frac{1}{2},\frac{1}{2}\rangle)$	$m_O(A,\langle 1,0,0\rangle)$	$m_O(A,\langle 0,0,0\rangle)$	$\sum$	Max	Gmax	Π^*	AM
$\langle 0,0,0\rangle$	1	1	0	2	1	(1,1,0)	4	0.66
$\langle \frac{1}{2},0,0\rangle$	0	0	0	0	0	(0,0,0)	1	0
$\langle 0,0,\frac{1}{2}\rangle$	1	1	0	2	1	(1,1,0)	4	0.66
$\langle \frac{1}{2},0,\frac{1}{2}\rangle$	0	0	0	0	0	(0,0,0)	1	0
$\langle 1,0,0\rangle$	0	0	1	1	1	(1,0,0)	2	0.33
$\langle 0,0,1\rangle$	1	2	1	4	2	(2,1,1)	12	1.33
$\langle 1,0,\frac{1}{2}\rangle$	0	0	1	1	1	(1,0,0)	2	0.33
$\langle \frac{1}{2},0,1\rangle$	0	1	1	2	1	(1,1,0)	4	0.66
$\langle 1,\frac{1}{2},1\rangle$	0	1	2	3	2	(2,1,0)	6	1
$\langle 0,\frac{1}{2},0\rangle$	1	1	0	2	1	(1,1,0)	4	0.66
$\langle \frac{1}{2},\frac{1}{2},0\rangle$	0	0	0	0	0	(0,0,0)	1	0
$\langle 0,\frac{1}{2},\frac{1}{2}\rangle$	1	1	0	2	1	(1,1,0)	4	0.66
$\langle \frac{1}{2},\frac{1}{2},\frac{1}{2}\rangle$	0	0	0	0	0	(0,0,0)	1	0
$\langle 1,\frac{1}{2},0\rangle$	0	0	1	1	1	(1,0,0)	2	0.33
$\langle 0,\frac{1}{2},1\rangle$	1	2	1	4	2	(2,1,1)	12	1.33
$\langle 1,\frac{1}{2},\frac{1}{2}\rangle$	0	0	1	1	1	(1,0,0)	2	0.33
$\langle \frac{1}{2},\frac{1}{2},1\rangle$	0	1	1	2	1	(1,1,0)	4	0.66
$\langle 1,1,1\rangle$	0	2	3	5	3	(3,2,0)	12	1.66

Table 3.4: The optimistic metric between the sequences in π and the elements of $\mathbb{A}$. Note that since the sequence in which all judgments are $\frac{1}{2}$ will always be closest to any judgment sequence, we can disregard it.

We show the inclusion properties of the distance-based aggregation rules built upon the specific aggregation functions we considered.

Proposition 3.4.2. $\Delta^{d,\Sigma} = \Delta^{d,AM}$ *for d such that* $\circledast = \sum$.

Proof. For $n \geqslant 1$

$$\sum_{i=1}^{n}\sum_{j=1}^{m} w(i,j)\cdot\delta(A\triangledown a_j, A_i\triangledown a_j) < \sum_{i=1}^{n}\sum_{j=1}^{m} w(i,j)\cdot\delta(A'\triangledown a_j, A_i\triangledown a_j)$$

if and only if

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m} w(i,j)\cdot\delta(A\triangledown a_j, A_i\triangledown a_j) < \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m} w(i,j)\cdot\delta(A'\triangledown a_j, A_i\triangledown a_j).$$

□

Proposition 3.4.3. $\Delta^{d,max} \subset \Delta^{d,Gmax}$ *and* $\Delta^{d,Gmax} \not\subset \Delta^{d,max}$ *for d such that* $\circledast = \sum$.

A	$d_H(A,\langle 1,0,0\rangle)$	$d_H(A,\langle 1,1,1\rangle)$	$d_H(A,\langle 0,0,0\rangle)$	Max	Gmax
$\langle 0,0,0\rangle$	1	3	0	3	(3,1,0)
$\langle 0,1,0\rangle$	2	2	1	2	(2,2,1)
$\langle 1,0,0\rangle$	0	2	1	2	(2,1,0)
$\langle 1,1,1\rangle$	2	0	3	3	(3,2,0)

Table 3.5: The Hamming metrics between the sequences in π and the elements of $\mathbb{A}$. The pink fields are the minima in the corresponding column.

Proof. For $x_1,\ldots,x_n,y_1,\ldots,y_n \in \mathbb{R}^+$ if $Gmax(x_1,\ldots,x_n) < Gmax(y_1,\ldots,y_n)$ then the first element of $Gmax(x_1,\ldots,x_n)$ is smaller or equal to the first element of $Gmax(y_1,\ldots,y_n)$. Since the first elements of $Gmax(x_1,\ldots,x_n)$ is $max(x_1,\ldots,x_n)$ and the first element of $Gmax(y_1,\ldots,y_n)$ is $max(y_1,\ldots,y_n)$. Consequently, if

$$Gmax_{i=1}^{n}\sum_{j=1}^{m} w(i,j)\cdot\delta(A\triangledown a_j, A_i\triangledown a_j) < Gmax_{i=1}^{n}\sum_{j=1}^{m} w(i,j)\cdot\delta(A'\triangledown a_j, A_i\triangledown a_j)$$

then

$$max_{i=1}^{n}\sum_{j=1}^{m} w(i,j)\cdot\delta(A\triangledown a_j, A_i\triangledown a_j) \leqslant max_{i=1}^{n}\sum_{j=1}^{m} w(i,j)\cdot\delta(A'\triangledown a_j, A_i\triangledown a_j).$$

To show that $\Delta^{d,Gmax} \not\subset \Delta^{d,max}$, it is sufficient to give an example of π. Consider $d = d_H$ and $W = U$. Let $\mathcal{A} = \{a_1,a_2,a_3\}$, $\mathcal{R} = \{a_3 \leftrightarrow a_1 \wedge a_2\}$ and

$$\pi = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

As it can be observed in Table 3.5, $\Delta^{d_H,max}(\pi,U) = \{\langle 1,0,0\rangle,\langle 0,1,0\rangle\}$, while $\Delta^{d_H,Gmax}(\pi,U) = \{\langle 1,0,0\rangle\}$. □

Proposition 3.4.4. $\Delta^{d,Gmax} \not\approx \Delta^{d,\Sigma}$ *for d such that* $\circledast = \sum$.

Proof. We give a counter-example.

Let $\mathcal{A} = \{a_1,a_2,a_3\}$, and $\mathbb{A} = \{\langle 0,0,0\rangle,\langle 0,1,1\rangle,\langle 1,0,0\rangle,\langle 1,1,0\rangle\}$, $d = d_H$, $W = U$ and

$$\pi = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

As it can be observed in Table 3.6, $\Delta^{d,Gmax}(\pi,U) = \{\langle 1,0,0\rangle\}$ while $\Delta^{d,max}(\pi,U) = \{\langle 1,1,0\rangle\}$.

A	$d_H(A,\langle 1,1,0\rangle)$	$d_H(A,\langle 1,1,0\rangle)$	$d_H(A,\langle 0,0,0\rangle)$	Gmax	$\sum$
$\langle 0,0,0\rangle$	2	2	0	(2,2,0)	4
$\langle 0,1,1\rangle$	2	2	2	(2,2,2)	6
$\langle 1,0,0\rangle$	1	1	1	(1,1,1)	3
$\langle 1,1,0\rangle$	0	0	2	(2,0,0)	2

Table 3.6: The Hamming metrics between the sequences in π and the elements of $\mathbb{A}$. The pink fields are the minima in the corresponding column.

□

Proposition 3.4.5. $\Delta^{d,\Sigma} \not\approx \Delta^{d,\Pi^*}$ *where d is such that* $\circledast = \sum$.

Proof. We give a counter-example.

Let $\mathcal{A} = \{a_1,a_2,a_3,a_4,a_5,a_6,a_7,a_8,a_9,a_{10},a_{11},a_{12},a_{13},a_{14}\}$ be an agenda. The set of all consistent judgment sets for it $\mathbb{A}$ is given in Table 3.7.

	$\mathcal{A} = \{$	$a_1,$	$a_2,$	$a_3,$	$a_4,$	$a_5,$	$a_6,$	$a_7,$	$a_8,$	$a_9,$	$a_{10},$	$a_{11},$	$a_{12},$	$a_{13},$	$a_{14}\}$
	$A^1 = \{$	$a_1,$	$\neg a_2,$	$\neg a_3,$	$a_4,$	$\neg a_5,$	$\neg a_6,$	$a_7,$	$\neg a_8,$	$\neg a_9,$	$a_{10},$	$\neg a_{11},$	$\neg a_{12},$	$a_{13},$	$a_{14}\}$
	$A^2 = \{$	$\neg a_1,$	$a_2,$	$\neg a_3,$	$\neg a_4,$	$a_5,$	$\neg a_6,$	$\neg a_7,$	$a_8,$	$\neg a_9,$	$\neg a_{10},$	$a_{11},$	$\neg a_{12},$	$a_{13},$	$a_{14}\}$
$\mathbb{A}$	$A^3 = \{$	$\neg a_1,$	$\neg a_2,$	$a_3,$	$\neg a_4,$	$\neg a_5,$	$a_6,$	$\neg a_7,$	$\neg a_8,$	$a_9,$	$\neg a_{10},$	$\neg a_{11},$	$a_{12},$	$a_{13},$	$a_{14}\}$
	$A^4 = \{$	$\neg a_1,$	$\neg a_2,$	$\neg a_3,$	$\neg a_4,$	$\neg a_5,$	$\neg a_6,$	$\neg a_7,$	$\neg a_8,$	$\neg a_9,$	$\neg a_{10},$	$\neg a_{11},$	$\neg a_{12},$	$\neg a_{13},$	$a_{14}\}$
	$A^5 = \{$	$\neg a_1,$	$\neg a_2,$	$\neg a_3,$	$\neg a_4,$	$\neg a_5,$	$a_6,$	$\neg a_7,$	$a_8,$	$\neg a_9,$	$\neg a_{10},$	$\neg a_{11},$	$\neg a_{12},$	$\neg a_{13},$	$\neg a_{14}\}$

Table 3.7: The set $\mathbb{A}$ of consistent judgment sets for agenda $\mathcal{A}$.

Let the profile π be such that $\pi \triangleright 1 = A_1$, $\pi \triangleright 2 = A_2$ and $\pi \triangleright 3 = A_3$. As it can be observed in Table 3.8, $\Delta^{d_H,\Sigma}(\pi,U) = \{A^1,A^2,A^3\}$, while $\Delta^{d_H,\Pi^*}(\pi,U) = \{A^4\}$.

□

Proposition 3.4.6. $\Delta^{d,Gmax} \not\approx \Delta^{d,\Pi^*}$ *for d such that* $\circledast = \sum$.

Proof. Consider the same example of π as in the proof of Proposition 3.4.5; $\Delta^{d_H,Gmax}(\pi,U) = \{A^1,A^2,A^3\}$, while $\Delta^{d_H,\Pi^*}(\pi,U) = \{A^4\}$. □

$A \in \mathbb{A}$	$d_H(A,A_1)$	$d_H(A,A_2)$	$d_H(A,A_3)$	$\sum$	Π^*
A^1	0	8	8	16	81
A^2	8	0	8	16	81
A^3	8	8	0	16	81
A^4	5	5	5	15	216
A^5	8	6	4	18	315

Table 3.8: The sum and product of Hamming metrics from an element in the set $\mathbb{A}$ to each of the agent's judgment sequences.

3.5 Representational abilities of $\Delta^{d,\odot}$

In this section we discuss the expressiveness of $\Delta^{d,\odot}$ in terms of judgment aggregation problems it can be applied to. We show how the co-domain can be controlled to obtain desirable properties for the collective judgment sequences. We also show how one can emulate the premise and conclusion based procedures using the $\Delta^{d,\odot}$ rules.

3.5.1 Co-domain restrictions for $\Delta^{d,\odot}$

A desirable property of an aggregation rule is to aggregate incomplete judgment sets but select a complete collective judgment set. This means we want a judgment aggregation rule that has incomplete judgment sets in its domain, but only complete judgment sets in its co-domain. The co-domain of the $\Delta^{d,\odot}$ rules is the set of all consistent judgment sequences $\mathbb{A}$. As a consequence, this property of completeness of the collective judgment sets is not satisfied by the $\Delta^{d,\odot}$ family. However we can extend the definition of $\Delta^{d,\odot}$ to include the co-domain as an additional parameter of the function.

Definition 43. *Let $X \subseteq \mathbb{A}$ be the subset of judgment sequences that satisfy a certain property. A X-restricted distance-based judgment aggregation rule is the rule $\Lambda^{d,\odot} : \mathbb{A}^n \times \mathbb{R}^{n\times m} \times \mathcal{P}(\mathbb{A}) \mapsto \mathcal{P}(\mathcal{P}(\mathbb{A}))$ defined as:*

$$\Lambda^{d,\odot}(\pi, W, X) = \underset{A\in X}{\arg\min} \odot_{i=1}^{n}(\circledast_{j=1}^{m} w_{i,j} \cdot \delta(A(j), p_{i,j})).$$

To ensure that the selected judgment sequences are complete, one needs to set $X = \mathbb{A}^{\downarrow 01}$. Restricting the co-domain can also be used to engineer that all collective judgment sets adhere to the view of the majority on particular agenda issues. As we know from the impossibility results in judgment aggregation such as (Dietrich, 2007; Pauly and van Hees, 2006), for most logics, the issue-majoritarian set is not always a consistent judgment set. However, for some subset of agenda issues $B \subset \mathcal{A}$, majority-adherence can be consistent and guaranteed. This subset B must be such that for all $a \in B$ and any valuation, $B\backslash\{a\} \cup \mathcal{R} \not\models_3 a$.

3.5.2 Emulating other judgment aggregation rules with $\Lambda^{d,\odot}$

The first two judgment aggregation "rules" are the premise-based and conclusion-based procedure presented in (Kornhauser and Sager, 1993), under the names "issue-by-issue voting and "case-by-case voting" respectively. These rules are applicable to agendas $\mathcal{A}$ that can be partitioned to a set of premises $\mathcal{A}^p$ and a set of conclusions $\mathcal{A}^c$.

Of the two procedures, the premise-based one has been more extensively studied in, for instance (Dietrich and Mongin, 2010; Mongin, 2008; Endriss et al., 2010b). According to the premise-based procedure, the collective judgment on a premise is the judgment supported by a majority of agents. According to the conclusion-based procedure, only the collective judgments on the conclusions are derived, by selecting that judgment for each conclusion that is supported by a majority of agents. Due to the "mandatory" incompleteness of the selected judgment sets, this procedure is not much considered in the literature. We extended the conclusion-based procedure with a distance-based procedure in (Pigozzi et al., 2009) to obtain collective judgments on the premises as well. Here we build on the work presented in (Pigozzi et al., 2009).

When the judgments are three-valued, there are two ways to define the majority function. One is the m_1 function which is used when aggregating binary profiles in Chapter 2.

Definition 44. *Let* $N_1 = \{i \mid \pi_{i,j} = 1\}$ *and* $N_0 = \{i \mid \pi_{i,j} = 0\}$. *The function* $m_1 : \mathbb{A}^n \times (\mathbb{R}^+)^{n\times m} \times \mathcal{A} \mapsto \{0,1\}$ *is defined as:*

$$m_1(\pi, a_j) = \begin{cases} 1 & \textit{iff } \sum_{i\in N_1} w_{i,j} > \sum_{i\in N_0} w_{i,j} \\ 0 & \textit{iff } \sum_{i\in N_1} w_{i,j} < \sum_{i\in N_0} w_{i,j} \\ \frac{1}{2} & \textit{iff otherwise} \end{cases}$$

The m_1 function undefined when $\sum_{i\in N_1} w_{i,j} = \sum_{i\in N_0} w_{i,j}$ and biased against the undecided judgment, namely the $\frac{1}{2}$ is only selected if the number of agents who render judgment 1 is the same as the number of agents who render the judgment 0. An unbiased majority function can be defined as well.

Definition 45. *Let* $N_1 = \{i \mid \pi_{i,j} = 1\}$, $N_{\frac{1}{2}} = \{i \mid \pi_{i,j} = \frac{1}{2}\}$ *and* $N_0 = \{i \mid \pi_{i,j} = 0\}$. *The unbiased, or absolute, majority function* $m_2 : \mathbb{A}^n \times (\mathbb{R}^+)^{n\times m} \times \mathcal{A} \mapsto \{0,1\}$ *defined for* $a \in \mathcal{A}$ *as:*

$$m_2(\pi, a_j) = \begin{cases} 1 & \textit{iff } \sum_{i\in N_1} w_{i,j} > \sum_{i\in N_0} w_{i,j} + \sum_{i\in N_{1/2}} w_{i,j} \\ 0 & \textit{iff } \sum_{i\in N_0} w_{i,j} > \sum_{i\in N_1} w_{i,j} + \sum_{i\in N_{1/2}} w_{i,j} \\ \frac{1}{2} & \textit{iff otherwise} \end{cases}$$

$Maj(\pi, W) = \langle m_2(\pi\triangledown 1, W), \ldots, m_2(\pi\triangledown m, W)\rangle$

The function m_2 is undefined when there is no one judgment that is supported by a majority of agents in a pair-wise compartment (with the other two judgments). For ternary judgment profiles we can define as many premise-based procedures as there are majority functions that can be defined.

Definition 46. *Given a profile* $\pi \in \mathbb{A}^n$, *an agenda* $\mathcal{A} = \{a_1^p, \ldots, a_k^p\} \cup \mathcal{A}^c$ *and* $\mathcal{R}$. *The biased premise-based procedure* $B-PBP$ *and the unbiased premise-based procedure* $U-PBP$ *is defined as*

$$B-PBP(P) = \{m_1(\pi, a^p) \mid a^p \in \mathcal{A}^p\} \cup \{a^c \mid a^c \in \overline{\mathcal{A}^c}, \{m_1(\pi, a_1^p), \ldots m_1(\pi, a_k^p)\} \cup \mathcal{R} \models_3 a^c\};$$

$$U-PBP(P) = \{m_2(\pi, a^p) \mid a^p \in \mathcal{A}^p\} \cup \{a^c \mid a^c \in \overline{\mathcal{A}^c}, \{m_2(\pi, a_1^p), \ldots m_2(\pi, a_k^p)\} \cup \mathcal{R} \models_3 a^c\}.$$

Another way to view the premise- and conclusion-based procedures is as rules in which the adherence to majority is guaranteed for the set of premises, or the set of conclusions correspondingly. We can represent and extend the premise- and conclusion-based procedures through a distance-based aggregation rule $\Lambda^{d,\odot}$.

The premise-based procedure is only applicable for those profiles P for which the set of premises is logically independent. The premise sub-profile π^p is the matrix obtained from the sub-sequences containing only judgments on the premises. Intuitively, π^p is the matrix obtained when from π by removing the columns corresponding to the elements of $\mathcal{A}^c$.

E.g., let agenda $\mathcal{A}$ be such that $\mathcal{A}^p = \{p, p \to q\}$ and $\mathcal{A}^p = \{q\}$. If π is a profile for $\mathcal{A}$, then π^p is the premise only sub-profile.

$$\pi = \begin{bmatrix} 1 & 0 & 0 \\ 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 1 \end{bmatrix} \qquad \pi^p = \begin{bmatrix} 1 & 0 \\ 1 & \frac{1}{2} \\ 0 & 1 \end{bmatrix}$$

We can extend the biased and unbiased premise-based procedures $B-PBP$ and $U-PBP$ to weighted distance-based judgment aggregation rules in the following way.

Definition 47 (Extended premise-based procedures)**.** *Let X_{bp} and X_{up} be co-domain restrictions defined as:*
X_{bp}: $A \in X_{bp}$ if and only if $A(a) = m_1(\pi, a)$ for all $a \in \mathcal{A}^p$
X_{up}: $A \in X_{up}$ if and only if $A(a) = m_2(\pi, a)$ for all $a \in \mathcal{A}^p$.

The biased and unbiased premise-based weighted aggregation functions are defined as:
$b - pbp(\pi, W) = \Lambda^{d,\odot}(\pi, W, X_{bp})$ and
$u - pbp(\pi, W) = \Lambda^{d,\odot}(\pi, W, X_{up})$.

If the agenda is such that the judgments on the conclusions are uniquely determined by the judgments on the premises, then the choice of d and $\odot$ are irrelevant. Otherwise there will be as many premise-based procedures as there are $(d, \odot)$ pairs.

In the similar manner we can define two extended conclusion-based procedures, $b-cbp$ and $u-cbp$.

Definition 48 (Extended conclusion-based procedures)**.** *Let the restrictions X_{bc} and X_{uc} be defined as X_{bc}: $A \in X_{bc}$ if and only if $A \downarrow a$ corresponds to $m_1(\pi, a)$ for all $a \in \mathcal{A}^c$*
X_{uc}: $A \in X_{uc}$ if and only if $A \downarrow a$ corresponds to $m_2(\pi, a)$ for all $a \in \mathcal{A}^c$.

We can define the extended conclusion-based procedures as:

$b - cbp(\pi, W) = \Lambda^{d,\odot}(\pi, W, X_{bc})$ and
$u - cbp(\pi, W) = \Lambda^{d,\odot}(\pi, W, X_{uc})$.

We can go about another way to extend the premise-based procedures, by using $\Delta^{d_T,\Sigma}$. Let $\mathcal{A}$ be an agenda containing only one issue a, $\mathcal{R} = \emptyset$ and π a profile for this agenda. We obtain that $\mathbb{A}(\mathcal{A}, \mathcal{R}, \models_3) = \{\langle 1 \rangle, \langle \frac{1}{2} \rangle, \langle 0 \rangle\}$.

$\sum_{i=1}^{n} |1 - A_i \nabla a| = |N_-| + |N_0|$,
$\sum_{i=1}^{n} |0 - A_i \nabla a| = |N_+| + |N_0|$ and
$\sum_{i=1}^{n} \lfloor |\frac{1}{2} - A_i \nabla a| \rfloor = |N_+| + |N_-|$.

The v selected by $\Delta^{d_T,\Sigma}(\pi,U)$ corresponds to the unbiased majority $m_2(\pi,a)$. When $\mathcal{A}$ has more issues, then $\Delta^{d_T,\Sigma}(\pi^p,U)$ returns the sequence corresponding to the set $\{m_2(P,a^p) \mid a^p \in \mathcal{A}^p\}$ because $\sum_{i=1}^{n}\delta_h(A\triangledown j,\pi i,j)$ are minimal for the judgments that correspond to the unbiased majority, due to

$$\begin{aligned} & \textstyle\sum_{i=1}^{n} d_T(A,A_i) \\ = & \textstyle\sum_{i=1}^{n}\sum_{j=1}^{m}\delta_T(A\triangledown j,\pi i,j) \\ = & \textstyle\sum_{j=1}^{m}\sum_{i=1}^{n}\delta_T(A\triangledown j,\pi i,j). \end{aligned}$$

Therefore, we can define an unbiased premise-based rule *UPBP* also as follows. Let $A^p = \Delta^{d_T,\Sigma}(\pi^p,U)$. We concatenate to the sequence A^p, the sequence of judgments on the conclusions obtained by deductively closing $\hat{A}^p$. We write

$$UPBP(\pi) \; = \; A^p \cup \langle\, v(a) \mid a \in \; A^c \text{ and } \hat{A}^p \; \cup \; \mathcal{R} \; \models_3 \; v(a)\rangle.$$

Another judgment aggregation rule frequently considered in the literature is the *sequential judgment aggregation procedure* (List, 2004a; Dietrich and List, 2007b; Li, 2010). This rule pre-supposes that there exists a total order $\geq$ over the agenda issues. The agenda issues are ranked according to some parameter as for example, relevance of the issue. The sequential procedure consists in applying the majority function m_1 to a subset of the agenda, starting from the issue ranked highest according to $\geq$ and continuing down the order until the judgments on the remaining issues are determined by the $m_1(\pi,a)$ already calculated. If the order is not total, then we can apply the $\Lambda^{d,\odot}$ rules to ensure that the collective judgments on the preferred issues correspond with $m_1(\pi,a)$. We can use m_2 as well.

3.6 Computational complexity of winner determination

Determining the complexity of the *winner determination problem* for social choice rules is one of the fields of research in the focus of computational social choice (Chevaleyre et al., 2007). The winner determination problem in voting theory is the problem of deciding whether a particular candidate is selected as the winner for a given profile of votes when a particular voting rule is used. The computational complexity of the winner determination problem is used as an indication of how difficult it is to determine the output from a particular social choice rule in the “worst case” profiles.

Endriss et al. (2010b) define the winner determination problem for judgment aggregation rules in terms of collective judgments instead of collective judgment sets. Given a number K, a profile π, an agenda $\mathcal{A}$ and a judgment for agenda issue a, the winer determination question is whether there exists a judgment sequence $A \in \mathbb{A}^{\downarrow 01}$ such that $A(a) = v(a)$ and the distance from A to the profile is smaller or equal than K. Endriss et al. (2010b) show that this winner determination problem, for $d = d_H$, $\odot = \sum$ and binary profiles, is solvable by a non-deterministic Turing machine in polynomial time.

In judgment aggregation, the winner determination problem can be defined as the problem of deciding whether a given judgment set is selected as the collective judgment set when a particular judgment aggregation rule is applied to a given profile of judgments. This is the approach we take because the distance based rules generate collective judgments for each issue. When the judgment aggregation rule is resolute, and the collective judgment sets are necessarily complete, it is a good choice to define the winner determination problem as in (Endriss et al., 2010b) since by checking for each judgment whether it is selected as collective or not, once can compute the entire collective judgment sets. However, when the judgment aggregation rule is irresolute, or the incomplete judgment sets can be selected as collective, this approach of checking judgment by judgment does not lead to computing a collective judgment sets.

We define the *judgment rule winner determination problem* and state it for an instance of an agenda $\mathcal{A}$, set of rules $\mathcal{R}$, and a judgment aggregation rule $F: \mathbb{A}^n \times (\mathbb{R}^+)^{n\times m} \mapsto \mathcal{P}(\mathbb{A})$ in the following way.

Definition 49 ($WinDet$ for F)**.**

Let F be a judgment aggregation rule $F: \mathbb{A}^n \times (\mathbb{R}^+)^{n\times m} \mapsto \mathcal{P}(\mathbb{A})$. We consider an agenda $\mathcal{A}$ of cardinality m, set of rules $\mathcal{R}$, and a set of agent n names N. The $WinDet$ problem for F is specified by the following input and output.

Input: *Profile $\pi \in (\mathbb{A}(\mathcal{A},\mathcal{R},\models_3))^n$, sequence $A \in \mathbb{A}(\mathcal{A},\mathcal{R},\models_3)$ and weight matrix $W \in (\mathbb{R}^+)^{n\times m}$.*
Output: *true if and only if $A \in F(\pi, W)$.*

We show the computational complexity of the $WinDet$ problem for $\Delta^{d,\odot}$ and $W = U$ without fixing the $\odot$ and d.

Proposition 3.6.1. *If $\odot$ and d are computable in polynomial time then $WinDet$ for $\Delta^{d,\odot}$ is in Σ_2^P.*

Proof. Σ_2^p is in the second level of polynomial-time hierarchy, see Figure 3.1.

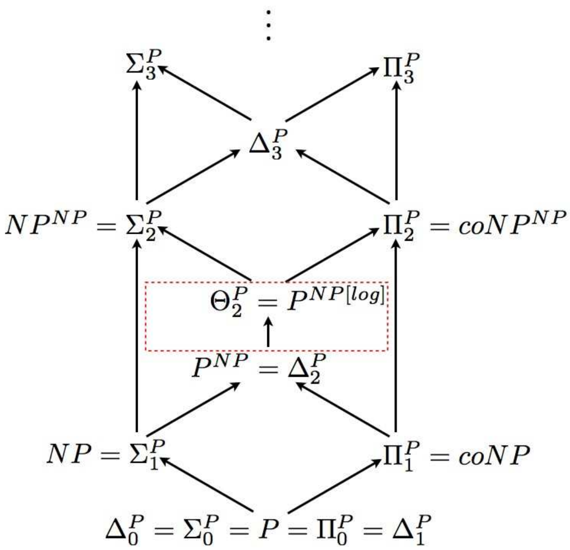


Figure 3.1: The polynomial time hierarchy, under the common assumption that **P**$\neq$**NP**. The arrows denote inclusion.

Σ_2^p is the class of all decision problems that can be solved by a non-deterministic Turing machine in polynomial time, *i.e.*, **NP** Turing machine, that has access to a non-deterministic oracle that takes polynomial time to respond to problems sent to it, *i.e.*, **NP** oracle (Papadimitriou, 1994, pg. 425). Each level i of the hierarchy is determined according to the following formulas, where the exponent is the class of the oracle:

- $\Delta_i^P = P^{\Sigma_{i-1}^P}$
- $\Sigma_i^P = NP^{\Sigma_{i-1}^P}$
- $\Pi_i^P = coNP^{\Sigma_{i-1}^P}$

We prove the proposition by showing an algorithm for $WinDet$ for $\Delta^{d,\odot}$.

Algorithm: $\underline{WinDet(\pi, A)}$

1. guess a valuation v for the atoms in $\mathcal{A}$;
2. if v is a model for A and not $ExistBetter(\pi, A)$
   then $return(true)$ else $return(false)$;

Oracle: $\underline{ExistBetter(\pi, A)}$

1. guess $A' \in \{0, \frac{1}{2}, 1\}^m$;
2. guess a valuation v' for the atoms in $\mathcal{A}$;
3. if v' is a model for A' and $\odot(d(A', \pi_1), \ldots, d(A', \pi_n)) > \odot(d(A, \pi_1), \ldots, d(A, \pi_n))$ then $return(true)$ else $return(false)$;

A distance from a judgment sequence A to a profile π is the output of $\odot_{i=1}^{n} d(A, \pi \triangleright i)$. The algorithm proceeds as follows. Recall that $\mathcal{A}$ can contain both atomic and non-atomic formulas. First, a valuation for the atoms in $\mathcal{A}$ is guessed. We check that the valuation is a model for A, *i.e.*, such that for its corresponding judgment set $\hat{A}$ it holds $\hat{A} \cup \mathcal{R} \not\models \bot$. Then we make a call to the **NP** oracle who returns a Boolean answer to the question of whether a judgment sequence exists that is closer to the profile than A. If such a sequence is found, A is not among the selected collective judgment sets.

The oracle determines its answer in the following way. First a judgment sequence A' is guessed and then a valuation for the atoms in A' is guessed. We ensure that A' is consistent and then we compare the distance from A' to the profile and the distance from A to the profile. The algorithms and the result can be easily adapted for the weighted case of $\Delta_W^{d,\odot}$.

□

The assumption we make for $\odot$ and d is that they are computable in polynomial time. The distances and aggregation functions we introduced in Section 3.3 are computable in polynomial time, with the possible exception of the distance d_G proposed in (Duddy and Piggins, 2011).

To calculate $d_G(A_1, A_2)$, one has to determine the shortest path between the vertices corresponding to A_1 and A_2 in the un-weighted bidirectional graph $\mathcal{G} = (\mathbb{A}^{\downarrow 01}, \mathcal{E})$. Finding the shortest path in a graph with non-negative weights can be solved using the Dijkstra's algorithm (Dijkstra, 1959) in quadratic time over the number of vertices. However, complexity is added by constructing the graph G. The set $\mathcal{E}$ is the set of edges defined as $(A_1, A_2) \in \mathcal{E}$ if and only if there exists no $A \in \mathbb{A}^{\downarrow 01}$ such that $d_H(A_1, A_2) = d_H(A_1, A) + d_H(A, A_2)$. Consequently, to construct G one has to check, for each A_1 and A_2 that there is no A between them. We conjecture that this problem is not solvable in polynomial time.

In the weighted case, a weight matrix W is also a part of the input. If $\circledast$ and δ are computable in polynomial time with respect to the size of π and W, then so is d, and the above result can be easily adapted.

Let us define a *score* of a $A \in \mathbb{A}$, with respect to a weight matrix W and a profile π to be the distance from A to π:

$$s(A,\pi,W) = \odot_{i=1}^{n} \circledast_{j=i}^{m} w_{i,j} \cdot \delta(A \triangledown a_j, \pi_{i,j}).$$

Depending on W and d, the number of possible scores can be known. For instance, for a weight matrix in which $w_{i,j} = 1$ for all i and j, and the Hamming distance, the number of possible scores is exactly the cardinality of the agenda m plus one. If the number of possible scores for $\Delta^{d,\odot}$ is known in advance and bounded by a polynomial in n,m then computing *WinDet* for $\Delta^{d,\odot}$ is in Θ_2^P. $\Theta_2^P = \mathbf{P}^{\mathbf{NP[log\ n]}}$ is the class of problems solvable by a polynomial-time deterministic Turing machine asking at most $\mathcal{O}(log\ n)$ adaptive queries to an **NP** oracle).[1]

The (conditional) membership in Θ_2^P can be demonstrated by the following variation of the algorithm.

For an ordered set X, let $med(X)$ denote the median of X, X^+ denote the subset of X from $med(X)$ up, and X^- the part below $med(X)$. Let *Val* be the set of possible scores.

Algorithm: $\underline{WinDet(\pi,A)}$

1. $Poss := Val$;
2. repeat
3. $\quad k := med(Poss)$;
4. $\quad$ if $Exist(\pi, Poss^-)$
   then $Poss := Poss^-$ else $Poss := Poss^+$;
5. until $|Poss| = 1$;
6. if $s(A,\pi,W) = med(Poss)$
   then $return(true)$ else $return(false)$;

Oracle: $\underline{Exist(\pi,Poss)}$

1. guess $A' \in \{0, \frac{1}{2}, 1\}^m$ and a valuation v;
2. if v is a model for A' and $s(A^{prime},\pi,W) \in Poss$ then $return(true)$ else $return(false)$;

According to this algorithm, we do a binary search to find the minimal score $s(A',\pi,W)$ that can be assigned to some $A' \in \mathbb{A}$. The binary-search algorithm can be executed in logarithmic number of steps, with respect to the size of *Val* and that is why there are logarithmic number of calls made to the oracle. If the score of the candidate A is this minimal score then A is among the sequences selected by $\Delta^{d,\odot}(\pi,W)$.

Endriss et al. (2010a) show that their winner determination problem for the premise-based procedure is decidable in polynomial time. The complexity does not change when we consider our *WinDet* for the $B-PBP(P)$ and $U-PBP(P)$ rules.

[1] I would like to thank an anonymous reviewer for the workshop of Social Choice and Artificial Intelligence held in conjunction with the 22nd International Joint Conference on Artificial Intelligence for hinting the property and sketching the proof.

Proposition 3.6.2. *The WinDet problem for $B-PBP(P)$ and $U-PBP(P)$ is in P.*

To prove this proposition, it is sufficient to give a description of an algorithm that decides if $\hat{A} \in B-PBP(P)$ or $\hat{A} \in B-PBP(P)$. First we check whether the candidate $\hat{A}$ is consistent. This is a model checking problem for ternary logic which can be solved in polynomial time. Then, for each judgment for a_j in $\hat{A}^p$ we check whether it corresponds to the biased or unbiased majority of $\pi^p \triangledown a_j$. This can be done in $O(m \cdot n)$ time.

The complexity of a winner determination problem only indicates how difficult it is to verify that a judgment sequence is selected. In judgment aggregation, it would be of interest to determine the complexity of the *search problem*. In a decision problem, we look for a confirmation whether a given output can be produced by a function. In a search problem, we instead look for the output that a function produces for a given argument.

The *WinDet* problem for $\Delta^{d_T,\Sigma}$ while $W=U$ is (still) in **NP**. This can be proved by slightly modifying the proof presented in (Endriss et al., 2010a).

Theorem 3.6.3. *The WinDet problem for $\Delta^{d_T,\Sigma}$ while $W=U$ is in* **NP** *for Kleene and Łukasiewitz logic.*

Proof. Endriss et al. (2010b) reduce their winner determination problem to the well known **NP** hard problem of integer programming, see for instance (Williams, 2009, Chapter 2). We can do the same. We write a variable $x_i \in \{0,1,2\}$ for each $a \in \mathcal{A}$ and $x_r \in \{0,1,2\}$ for each element of $r \in \mathcal{R}$. The constraints for x_r is $x_r > 0$ for each j. The constraints for x_i depend on the ternary logic used. We show the constraints for the Kleene and Łukasiewitz logic. The Kleene logic observes the relations between the connectives so we can give only the constraints for $\neg$ and $\wedge$:

$$\begin{aligned} a_2 = \neg a_1 : &\quad x_2 = 2 - x_1 \\ a_3 = a_1 \wedge a_2 : &\quad x_3 \leqslant x_1, x_3 \leqslant x_2, x_1 + x_2 \leqslant x_3 + 2 \\ a_3 = a_1 \leftrightarrow a_2 : &\quad x_1 + x_2 \leqslant x_3 + 2, x_1 + a_3 \leqslant x_2 + 2 \\ &\quad x_r > 0 \end{aligned} \tag{3.1}$$

For the Łukasiewitz logic $\varphi \to \varphi \equiv \neg(\neg\varphi \wedge \varphi)$ does not hold, hence we need to specify the constraints for x_i for all of the connectives:

$$\begin{aligned} a_2 = \neg a_1 : &\quad x_2 = 2 - x_1 \\ a_3 = a_1 \wedge a_2 : &\quad x_3 \leqslant x_1, x_3 \leqslant x_2, x_1 + x_2 \leqslant x_3 + 2 \\ a_3 = a_1 \vee a_2 : &\quad x_3 \geqslant x_1, x_3 \geqslant x_2, x_1 + x_2 \geqslant x_3 + 2 \\ a_3 = a_1 \to a_2 : &\quad x_3 \leqslant 2, x_3 < 2 - x_1 + x_2, x_1 + x_2 \leqslant x_3 + 2 \\ a_3 = a_1 \leftrightarrow a_2 : &\quad x_1 + x_2 \leqslant x_3 + 2, x_1 + a_3 \leqslant x_2 + 2, x_2 + x_3 \leqslant x_1 + 2, 2 \leqslant x_1 + a_2 + x_3 \\ &\quad x_r > 0 \end{aligned} \tag{3.2}$$

We omit the requirement in the proof of (Endriss et al., 2010b) that a for particular judgment a_j, the constraint $x_j = 1$ is added because we are not interested if one particular judgment is selected as collective or not. We set the score K to be the score of any of the sequences in the profile. The rest of the proof can be used unchanged.

Let $\pi \in \mathbb{A}^n$, $W = U$, and $A \in \mathbb{A}$. Let

$$n_1(A(j), \pi \triangledown j) = \#\{i \mid \delta(A(j) - \pi_{i,j}) = \frac{1}{2}$$
$$n_2(A(j), \pi \triangledown j) = \#\{i \mid \delta(A(j) - \pi_{i,j}) = 1 \tag{3.3}$$

For $d = d_T$ it holds that

$$\begin{aligned}
\sum_{i=1}^{n} d_T(A, \pi \triangleright i) &= \sum_{i=1}^{n}\sum_{j=1}^{m} \delta_T(A(j), \pi_{i,j}) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} |A(j) - \pi_{i,j}| \\
&= \sum_{j=1}^{m} (\cdot n_1(A(j), \pi \triangledown j) + 2n_2(A(j), \pi \triangledown j))
\end{aligned} \tag{3.4}$$

For $d = d_H$ it holds that

$$\begin{aligned}
\sum_{i=1}^{n} d_H(A, \pi \triangleright i) &= \sum_{i=1}^{n}\sum_{j=1}^{m} \delta_H(A(j), \pi_{i,j}) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} ||A(j) - \pi_{i,j}|| \\
&= \sum_{j=1}^{m} n_2(A(j), \pi \triangledown j))
\end{aligned} \tag{3.5}$$

To compute a winner under $\Delta^{d_T,\Sigma}$ we need to find a sequence $A \in \mathbb{A}(\mathcal{A}, \mathcal{R}, \models_{3L})$, or $A \in \mathbb{A}(\mathcal{A}, \mathcal{R}, \models_{3K})$ correspondingly, characterized by variables $x_1, \ldots, x_m$ that minimizes the sum

$$n_1(x_1, \pi \triangledown 1) + 2 \cdot n_2(x_1, \pi \triangledown 1) + \cdots + n_1(x_m, \pi \triangledown m) + 2 \cdot n_2(x_m, \pi \triangledown m).$$

For the case d_T we need to minimize the sum

$$n_2(x_1, \pi \triangledown 1) + \cdots + n_2(x_m, \pi \triangledown m).$$

To this end we introduce an additional set of integer variables $y_j > 0$ for $j \in [1, m]$. We ensure that $y_j = n_1(x_j, \pi \triangledown j) + 2 \cdot n_2(x_j, \pi \triangledown j)$ by adding the constraints :

$$\begin{aligned}
&(\forall j \leqslant m) \quad n_1(x_1, \pi \triangledown 1) + 2 \cdot n_2(x_j, \pi \triangledown j) \leqslant y_j \\
&(\forall j \leqslant m) \quad n_1(x_1, \pi \triangledown 1) + 2 \cdot n_2(x_j, \pi \triangledown j) \geqslant y_j
\end{aligned} \tag{3.6}$$

or the constraints

$$\begin{aligned}
&(\forall j \leqslant m) \quad n_1(x_1, \pi \triangledown 1) + 2 \cdot n_2(x_j, \pi \triangledown j) \leqslant y_j \\
&(\forall j \leqslant m) \quad n_1(x_1, \pi \triangledown 1) + 2 \cdot n_2(x_j, \pi \triangledown j) \geqslant y_j
\end{aligned} \tag{3.7}$$

correspondingly.

Now we need minimize

$$\sum_{j=1}^{m} y_j \leqslant K$$

subject to constraints (3.1) and (3.6) for Kleene logic and d_T, to constraints (3.1) and (3.7) for Kleene logic and d_H, to constraints (3.2) and (3.6) for Łukasiewicz logic and d_T, and to constraints (3.2) and (3.7) for Łukasiewicz logic and d_H. This integer program is feasible if and only if A is a winner for $\Delta d, \sum(\pi, U)$.

□

3.7 Conclusions

Judgment aggregation rules used by hierarchical groups should be able to aggregate incomplete judgment sets into complete judgment sets regardless of the number of agents or type of agenda. The rules should also be able to aggregate weighted judgments. In this chapter we develop a family of weight-sensitive distance-based judgment aggregation rules that satisfy these requirements.

A distance-based judgment aggregation rule is fully specified by specifying a pair $(d, \odot)$ of product metric d (specified by another aggregation function $\circledast$ and a metric δ) and aggregation function $\odot$. We present examples of distance functions and aggregation functions. While the Hamming and Drastic distances have already been used in judgment aggregation, the Taxicab distance is a new option. From the five aggregation functions we presented, $\sum$, *max* and *Gmax* have been already introduced in the literature of belief merging by (Konieczny and Pino-Pérez, 1999), but *AM* and Π^* are new. When applied to the same profile, *AM* gives the same results as $\sum$. The rule Π^*, however, gives rise to truly new aggregation operators. We summarize the (non) inclusion results between the families of distance-based aggregation rules in Table 3.9.

	$\Delta^{d,\sum}$	$\Delta^{d,AM}$	$\Delta^{d,max}$	$\Delta^{d,Gmax}$	Δ^{d,Π^*}
$\Delta^{d,\sum}$	$=$	$\subset$	$\neq$	$\neq$	$\neq$
$\Delta^{d,AM}$	$=$	$=$	$\neq$	$\neq$	$\neq$
$\Delta^{d,max}$	$\neq$	$\neq$	$=$	$\subset$	$\neq$
$\Delta^{d,Gmax}$	$\neq$	$\neq$	$\supset$	$=$	$\neq$

Table 3.9: The summary of the (non)inclusion results for d being a product metric constructed using $\circledast = \sum$.

In this chapter we also discuss the computational complexity of the winner determination problem for the distance-based aggregation rule. For an unspecified d and $\odot$ the complexity is Σ_2^P, which can be considered as high since Σ_2^P in the second level of the polynomial time hierarchy (Papadimitriou, 1994, pg. 425). This complexity is lowered to Θ_2^P when certain distances and aggregation operators are used, such as d_H and Σ. In Section 3.5.2 we introduced the extended premise-based procedure. The complexity of the winner determination problem for this procedure is considerably lower, but the premise-based procedure is not applicable to all agendas. Endriss et al. (2010a) show that checking whether an agenda is "safe" is a problem in the complexity class Π_2^P, which is in the same level in the hierarchy as Σ_2^P. Consequently, checking if the premise-based procedure can be applied safely is more difficult than using the procedure (Endriss et al., 2010a). The distance-based aggregation rules have a higher complexity, but they can always be applied.

Based on the complexity analysis in Section 3.6 we can make the conclusion that extending a distance-based rule from binary to ternary judgments does not influence the complexity of the *WinDet* problem for the rule. However, extending the rule from unweighted to weighted judgments can influence the *WinDet* complexity.

Compared to Chapter 2, here we do not generate as many specific rules. However, we do give a "template" for many weighted distance-based rules, each specified by a pair of aggregation functions and a metric. Therefore we need to be able to distinguish among all possible specific rules that can be generated using this "template". This means that we need to know how to select the aggregation functions $\odot$ and $\circledast$, and a metric δ so that the resulting rule is adequate for a given decision-reaching problem.

In part we answer the question of selecting $\odot$, $\circledast$ and δ by the complexity analysis of the *WinDet* problem. For well-behaved rule in terms of complexity of the *WinDet* problem, $\odot$, $\circledast$ and δ should be computable in polynomial time and the co-domains of both $\odot$ and $\circledast$ should be enumerable. However this characterization of $\odot$, $\circledast$ and δ is still very general. In the next chapter we define various structural and relational properties that enable us to further specify $\odot$, $\circledast$ and δ and distinguish further among weighted distance-based rules.

Part II

Selecting judgment aggregation rules

4

Selecting judgment aggregation rules

Abstract. We need a way to distinguish between the judgment aggregation rules we constructed. To accomplish this we qualify the judgment aggregation rules by the properties they satisfy. In the judgment aggregation literature such properties have been studied from a combinatorial, which properties are mutually consistent, or characterization point of view. One typically studies which minimal set of properties can be satisfied by a judgment rule. Alternatively one studies the properties that characterize all rules that select collective judgment sets from a desirable set of judgment sets. Compared with properties studied in voting theory, not many properties in judgment aggregation have been considered for judgment aggregation rules. In this chapter we construct properties for judgment aggregation rules and we study which of our rules satisfy them.

4.1 Introduction

In Chapters 2 and 3 we introduced many judgment aggregation rules producing distinct judgment sets for the same profile. How can we choose which rule to use for a given multi-agent system group decision problem? A judgment aggregation rule is a function that assigns a non-empty set of collective judgment sets to a profile of individual judgments. How good is a specific judgment aggregation rule for a particular problem?

The conventional approach to qualifying aggregation rules in social choice theory is a theoretical analysis of properties; one conceptualizes, defines and studies (un)desirable properties for the rules. In this chapter we take the theoretical approach to qualifying judgment aggregation rules. In social choice theory, one studies which minimal set of properties can be satisfied at the same time by a non-dictatorial or non-oligarchic judgment aggregation rule, such is the work of (Dietrich and List, 2008a; Nehring and Puppe, 2010b), or the properties as a way to characterize the rules that select from a specific collection of judgment sets, such as (Grandi and Endriss, 2010, 2011; Nehring et al., 2011; Nehring and Pivato, 2011). This gives rise to impossibility results along the line of the Arrow's theorem (Arrow, 1963, Chapter 3).

In this chapter we are interested in constructing desirable properties for judgment aggregation rule and in analyzing which of the rules introduced in Chapters 2 and 3 satisfy the constructed properties. Each aggregation context gives rise to its own set of necessary, desirable, and undesirable properties for a rule.

The properties of a judgment aggregation rule can be classified in two large groups: *structural characteristics* and *relational properties*. The structural characteristics are the properties that are satisfied by the profile being aggregated and by the collective sets that are the aggregate, *i.e.*, these are properties of the domain and co-domain of the judgment aggregation rule. The relational properties are the properties that hold between the input profile and each of the assigned judgment sets. It is mainly the relational properties that have been considered in the judgment aggregation literature.

The first desirable properties considered for judgment aggregation are universal domain, anonymity and systematicity (List and Pettit, 2002), directly "imported" from the preference aggregation conditions of Arrow (Arrow, 1963, Chapter 3). Further properties have been considered: independence of irrelevant information (Dietrich, 2006a), monotonicity properties (Nehring and Puppe, 2010a; Dietrich and List, 2005; List and Puppe, 2009), and (Dietrich and List, 2008a) as well as unanimity properties by for instance (Dietrich and List, 2008b; List and Puppe, 2009).

In the judgment aggregation literature (List and Polak, 2010), but also in the abstract and binary aggregation literature (Dokow and Holzman, 2010a; Grandi and Endriss, 2011), one considers a judgment aggregation *function* to be a function which associates a profile of binary judgments to a unique complete set of binary judgments, Figure 4.1 a). The listed properties, universal domain, anonymity, systematicity, the monotonicity and the unanimity properties, are defined for such judgment aggregation functions. In Chapter 2 we defined judgment aggregation rules which are functions that associate a profile of binary judgments to *set of a possibly incomplete judgment sets*, Figure 4.1 b). In Chapter 3 the judgment aggregation rules are extended to functions that associate a pair of profiles, of ternary judgments and associated weights, to a set of sequences of ternary judgments, Figure 4.1 c).

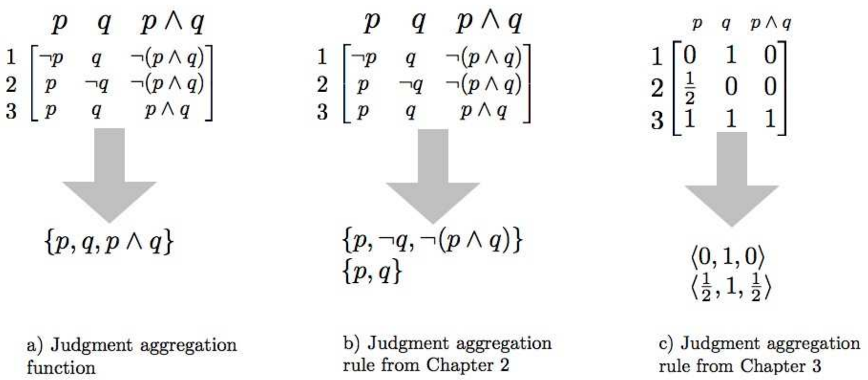


Figure 4.1: Illustration of the different ways to define judgment aggregation functions.

The structural properties of judgment aggregation, particularly those referring to the domain, can be considered without drastically adapting the definition available in the literature. However, for studying relational properties one needs to construct definitions for judgment aggregation rules that correspond to the properties defined for judgment aggregation functions. The main culprit for these difficulties is the *irresoluteness* of judgment aggregation rules. A judgment aggregation rule is *resolute* when for every profile, and every weight matrix in the case of a weighted rule, the aggregate is a singleton set; and irresolute otherwise.

Consider for instance the unanimity principle which states that if all agent rendered the same judgment for issue $a \in \mathcal{A}$ then this is the judgment for a in the collective judgment set. Since we have more collective judgment sets, is the unanimity principle satisfied if the unanimous judgment is contained in at least one, most of, or all of the collective judgment sets? We need to define the relational properties for the generalized irresolute version of judgment aggregation rules. We construct these definitions by first defining when does a property of a resolute rule corresponds to a property of an irresolute rule. As hinted by the example of the unanimity property, there can be several versions of an irresolute rule properties that correspond to a given resolute rule property. We consider the most common properties encountered in the judgment aggregation literature.

A common concern when constructing a social choice rule is its response to manipulative agents. An agent is manipulative if he choses the information to submit (opinions, judgments, preferences, votes, etc.) instead of being honest, with the purpose of ensuring that the aggregate is one he prefers. A social choice rule is *strategy-proof* when none of the agents can benefit by being manipulative. Manipulability of judgment aggregation rules is more difficult to study compared to that of *e.g.*, voting theory, since the agents are not modeled to hold preferences over the judgment sets. The research on manipulability of judgment aggregation function, by Dietrich and List (2005) and Endriss et al. (2010b), is conducted by making the assumption that the smaller the distance between an agent's judgment set and a collective judgment set, the more that agent "prefers" that collective judgment set. The same approach is taken by Everaere et al. (2007) for studying the manipulability of distance-based belief merging operators.

Apart from the properties considered in the judgment aggregation literature, there are many interesting properties that can be imported from voting theory. These are of interest for judgment aggregation rules as well, both from social-theoretic and computational viewpoint. In particular we can outline the property of *separability* (Smith, 1973), also called *consistency* by (Young, 1975) and the *independence of clones* property (Tideman, 1987). The property of separability in voting theory states that, given two profiles of votes for the same set of candidates, if a candidate is among the winner of the both profiles, then the same candidate is among the winners of the profile obtained by combining the both profiles. This property is desirable from a computational point of view. Aggregating smaller profiles of judgments is more efficient than aggregating large profiles. If a rule that satisfies separability is used, a very large group of agents can be split and their profiles aggregated separately. Comparing the winners from each profile can eliminate the need to aggregate the joint profile of all the agents.

The property of independence of clones in voting theory states that when a candidate is added to the set of candidates, and this candidate is identical to a candidate already in the set, the winner of the election does not change. There are no candidates in judgment aggregation, but the independence of clones can be defined for agenda items. Indeed, if two agenda issues are logically equivalent, then the collective judgment on the both equivalent issues should be the same, regardless the profile. We define the separability and independence of clones properties in a judgment aggregation framework.

The manipulability of a social choice rule is an important issue when the outcome of the rule constrains in some way the behavior of the agents. For example, a person in a group of friends, choosing a restaurant for dinner by voting on a selection of restaurants, is expected by social norm to go to the restaurant that is selected by the voting rule, even if he does not particularly prefer it. In an agreement reaching setting, an agent is constrained to abide by the results of the judgment aggregation, *e.g.*, participate in solving the problem once a solution is agreed on. However, we made the assumption that the agents are non-manipulative and thus we are not concerned with this aspect of the judgment aggregation rules.

In some aggregation contexts it is possible to evaluate the performance of a judgment aggregation rule experimentally. Consider Example 4.1.1.

Example 4.1.1. *A group of three agents needs to agree on whether there is a fire in the building they occupy. They consider the following set of issues:*

- *fire is observed* (f)*,*
- *smoke is observed* (s)*,*
- *the alarm went off* (a)*.*

The agenda is $\mathcal{A} = \{f, s, a\}$ *and the constraints are* $\mathcal{R} = \{s \to f, (a \wedge s) \to f\}$*. The robots interpret their perceptual data to construct an opinion regarding the truth-values of* f*,* s *and* a*. Since the perceptual data can be inconclusive or even wrong due to sensor malfunction, the robots may observe an* $a \in \mathcal{A}$ *as true even when* a *is false, or observe* a *as false even when it is true. However, regardless of what the robots observe, there is a unique factual truth-value assignment of* f*,* s *and* a *that corresponds to the actual state of the world.*

Truth-tracking is the process of establishing these unique truth-values that correspond to the actual state of the world (Hartmann et al., 2010). How good a judgment aggregation rule is can be measured with respect to how good the rule is at truth-tracking, namely how often it is the case that the collective judgment set assigns values that correspond to the actual state of the world.

If the aggregation context is such that truth-tracking is the goal of the collective decision then a good judgment aggregation rule is one that is good at truth-tracking. To ascertain the truth-tracking quality of a rule, one needs to construct a simulation or experiment in which one can compare the number of total aggregations with the number of aggregations in which the collective judgment set produced by said rule are faithful to the truth. The experimental approach to qualifying judgment aggregation rules in terms of truth-tracking is addressed in (Ganesan, 2011).

This chapter is structured as follows. In Section 4.2 we give a definition of an aggregator rule that generalizes the aggregators considered in Chapters 2 and 3. We also define when a property defined for one type of rule corresponds to a property defined for another type of rule. In Section 4.3 we define and analyze the structural properties considered in the judgment aggregation literature. In Sections 4.4 and 4.5 we define and analyze the first relational properties studied in judgment aggregation: independence of irrelevant information, neutrality and anonymity. In Section 4.6 we define and study majority-adherence properties, in Section 4.7 unanimity adherence properties, in Section 4.8 monotonicity properties and in Section 4.9 separability properties. In Section 4.10 we give the definitions of properties that can be desirable for a judgment aggregation rule, but we do not study how these properties are satisfied by the rules we considered in Chapters 2 and 3. Lastly in Section 4.11 we give an overview of which properties are satisfied by which rules and relate aggregation contexts with structural and relational properties that are advantageous for rules used in these contexts.

4.2 Preliminaries

We begin by constructing a general definition of a judgment aggregation rule. The rules introduced in Chapters 2 and 3, but also judgment aggregation functions in other literature, *e.g.*, (List and Polak, 2010; Dokow and Holzman, 2010b; Grandi and Endriss, 2011), are a special case of this general rule. Using this definition we can define properties of rules that are applicable for all categories of rules.

Definition 50. *Let N be a set of n agent names, T a finite enumerable set of truth-values, Ł a T-valued logic with an entailment operator $\models_T$, $\mathcal{A} \subseteq \mathcal{L}$ an agenda of m elements and $\mathcal{R} \subset \mathcal{L}$ a set of constraints. $\mathbb{A}(\mathcal{A},\mathcal{R},\models_L)$ is the set of all sequences from $T^{n\times m}$ that satisfy the constraints $\mathcal{R}$ and are consistent with respect to $\models_T$. Let $\mathbb{R}^+$ be the interval of reals $[0,+\infty)$ and $S_I, S_{II} \subseteq \mathbb{A}(\mathcal{A},\mathcal{R},\models_L)$. A weighted judgment aggregation rule is a function*

$$F : S_I^n \times (\mathbb{R}^+)^{n\times m} \mapsto \mathcal{P}(S_{II}).$$

We have so far worked with three types of sets S_I and S_{II}.

Example 4.2.1. *Let the set S_2 be the set of all consistent and complete (binary) judgment sequences. Namely, $S_2 = \mathbb{A}(\mathcal{A},\mathcal{R},\models) \subset \{0,1\}^m$, where $\models$ is the classical propositional logic entailment operator. A judgment aggregation function (binary aggregation rule, abstract aggregation rule) (List and Polak, 2010; Dokow and Holzman, 2010b; Grandi and Endriss, 2011) can be defined as*

$$f : S_2^n \times \{1\}^{n\times m} \mapsto S_2.$$

Example 4.2.2. *Let $S_3 = \mathbb{A}(\mathcal{A},\mathcal{R},\models_3) \subset \{0,\frac{1}{2},1\}^m$, where $\models_3$ is a classical entailment operator for some three-valued logic. The rules in Chapter 3 can be defined as*

$$\Delta^{d,\odot} : S_3^n \times (\mathbb{R}^+)^{n\times m} \mapsto \mathcal{P}(S_3).$$

Example 4.2.3. *Let $S_L = \mathbb{A}(\mathcal{A},\mathcal{R},\models_{3L}) \subset \{0,\frac{1}{2},1\}^m$ be the set of all ternary judgment sequences consistent with respect to the ternary Łukasiewicz logic semantics (Łukasiewicz, 1920; Urquhart, 2001). We can define the rules based on minimization in Chapter 2 trough Definition50 using S_L as a co-domain.*

Recall that most of these rules in Chapter 2 are such that the selected collective judgment sets are incomplete. Each judgment set A incomplete on a judgment for $a \in \mathcal{A}$ can be replaced with two complete judgment sets A^+ and A^- such that $a \in A^+$ and $\neg a \in A^-$; the resulting judgment sets being consistent under ternary Łukasiewicz logic semantics.

According to the semantics of the Łukasiewicz ternary logic, one can think of the truth-values as sets of classical truth-values, namely $0 = \{F\}$, $1 = \{T\}$ and $\frac{1}{2} = \{T,F\}$. The third value is in a sense interpreted as a variable that can be replaced with either true or false. Therefore, the co-domain of the aggregation rules from Chapter 2 is equivalent to $\mathcal{P}(S_L)$. The rules in Chapter 2, can be defined as

$$R : S_2^n \times \{1\}^{n\times m} \mapsto \mathcal{P}(S_L).$$

Observe that $S_2, S_L \subseteq \mathbb{A}(\mathcal{A},\mathcal{R},\models_{3L})$.

The relational properties of judgment aggregation rules that can be encountered in the literature are defined for functions that are resolute. How can a relational property defined for resolute rules be “lifted” to irresolute rules? Moreover, if we have a definition of relational property for irresolute rules, how does this property relate to a property defined for resolute rules?

To answer this question we introduce the concept of *lifting* between properties. Observe that, when restricted to the (possibly empty) domain $\mathbb{A}_1^n = \{\pi \mid \pi \in \mathbb{A}_c^n$ and $|F(\pi, U)| = 1\}$, F is a resolute rule. Recall that U was the weight matrix in which all weights $w(i, j) = 1$. We consider the unweighted case $W = U$ only since the resolute rules considered in the literature, and the properties defined for them, are unweight.

In addition to the difference in output cardinality, the resolute rules defined in the literature are defined for binary judgments, while the rules defined in Chapter 3 are defined for ternary judgments. Therefore, when constructing a definition of lifting, we also need to consider the relation between the sets S for which the aggregators are defined.

Definition 51. *Let the function $f : S^n \mapsto S$ be a resolute rule and let (γ) be a property for such defined functions. Let $F : S_I^n \times (\mathbb{R}^+)^{n\times m} \mapsto \mathcal{P}(S_{II})$ be a judgment aggregation rule that satisfies some property (χ). The sets S, S_I and S_{II} are such that $S \subseteq S_I$ and $S \subseteq S_{II}$. A property (χ)* lifts *a property (γ) if (γ) is satisfied by F for every $\pi \in S^n$ such that $|F(\pi, U)| = 1$.*

For a relational property, there will always be more than one possible (χ) property that lifts the same (γ) property due to the irresoluteness of the judgment aggregation rules. In the next chapters we construct properties by lifting the most common properties for resolute rules.

4.3 Structural properties

Once we have the definition of a rule, we can consider its structural properties. The type of judgments being aggregated characterizes the input, *i.e.*, the domain, of a judgment aggregation rule. We can distinguish between two orthogonal types: the *value-type* and the *weight-type*. The value-type specifies the values that the judgments in the profile can take. We have considered binary and ternary judgment profiles. Multi-valued profiles are considered in (Pauly and van Hees, 2006) and (Li, 2010). The weight-type specifies the weights that can be associated with the profile. We considered unweight profiles, agent-associated weights, agenda issue associated weights and judgment weights.

One structural property considered in the judgment aggregation literature is *universal domain*, defined in (List and Pettit, 2002) for judgment aggregation functions. The universal domain is satisfied when the judgment aggregation function is defined for all profiles of complete and consistent judgment sets. Universal domain is simple to generalize, since it only refers to the rule's domain. All introduced judgment aggregation rules satisfy universal domain by construction.

Definition 52. *A judgment aggregation rule F from Definition 50 satisfies the* universal domain *if and only if, for every $\mathcal{A}, \mathcal{R}$ and $\models_L$, $S_I \subseteq \mathbb{A}(\mathcal{A}, \mathcal{R}, \models_L)$.*

Collective rationality (List and Puppe, 2009) is another structural property considered in the literature. Collective rationality states that only rational collective judgments are admissible as outputs. This means that the judgment aggregation rule always selects consistent judgment sets. This property is also easy, when compared to other properties, to define for the judgment aggregation rules of Definition 50.

Definition 53. *A judgment aggregation rule F from Definition 50 satisfies* collective rationality *if and only if, for every $\pi \in S_I^N$ and $W \in (\mathbb{R}^+)^{n\times m}$, $F(\pi, W) \subseteq \mathbb{A}(\mathcal{A}, \mathcal{R}, \models_L)$.*

The judgment aggregation rules of Definition 50 satisfy collective rationality by construction.

The output, *i.e.*, the co-domain, of a judgment aggregation rule is also characterized by value-type as well, *e.g.*, binary in the case of f, Łukasiewicz ternary in the case of R, ternary in the case of Δ. The value-type of the output is expressed trough a property called *completeness* (List and Puppe, 2009). A judgment aggregation function satisfies completeness if it always selects complete judgment sets. We give a formal definition for rules.

Definition 54. *Let $\mathbb{A}^{\downarrow 01}$ be the binary restriction of the set $\mathbb{A}(\mathcal{A},\mathcal{R},\models_L)$. A judgment aggregation rule F from Definition 50 satisfies* completeness *if and only if, for every $\pi \in S_I^N$ and $W \in (\mathbb{R}^+)^{n\times m}$, $F(\pi,W) \subseteq \mathbb{A}^{\downarrow 01}$.*

Most of the rules we defined do not satisfy completeness. The exceptions are the rules R_{MWA} and $R^{d_H,max}$ in Chapter 2.

Additionally, the output is characterized by the cardinality, namely how many sequences are included in the output. According to cardinality, we distinguish between resolute rules that are also referred to as functions and always select a unique judgment sequence, and irresolute rules. All the rules we introduced in Chapters 2 and 3 are irresolute by construction. As multiple impossibility results in judgment aggregation show, resolute rules that satisfy some minimal desirable conditions can only be defined for restricted domains (List and Polak, 2010; Dietrich and List, 2010). We refer to resolute rules as *function aggregators* and irresolute rules as *rule aggregators*.

4.4 Independence of irrelevant information

One of the first properties considered in judgment aggregation (List and Pettit, 2002) is the property of systematicity. An unweight function aggregator for binary judgments satisfies systematicity if it satisfies *independence of irrelevant information*[1] and *neutrality*. The properties of systematicity and in particular independence of irrelevant information have been among the most debated in the judgment aggregation literature, with (Dietrich and List, 2005; Nehring and Puppe, 2005) deeming this property desirable, and (Chapman, 2002) discussing its controversies.

In voting theory, the property of neutrality states that the order of the candidates in the candidate set has no bearing on who is selected as the winner. In judgment aggregation, if it is to be a counterpart of the one in voting theory, then the property of neutrality should state that the collective judgment selected for a collection of individual judgments does not depend on the particular issue for which those judgment are rendered, *i.e.*, the order in which the issues in the agenda are given, does not influence the collective judgment set obtained. In judgment aggregation neutrality, when considered together with independence of irrelevant information, is taken to mean that each issue is aggregated using the same aggregation rule for each issue. Indeed, this latter view on neutrality is a consequence of the voting counterpart neutrality in the presence of independence of irrelevant information.

We illustrate the voting counterpart neutrality with an example. Consider the following two profiles for some agenda $\mathcal{A} = \{a_1,a_2,a_3\}$ and $\mathcal{A}' = \{a_2,a_1,a_3\}$ correspondingly:

[1] Independence of irrelevant information is also sometimes called independence of irrelevant alternatives.

$$\pi_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \qquad \pi_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

Observe that the judgments for a_1 in π_1 are the same as the judgments for a_2 in π_2. Assume that a rule f is applied to both π_1 and π_2, $A_1 = f(\pi_1)$ and $A_2 = f(\pi_2)$. If f is neutral, then necessarily $A_1(1) = A_2(2)$, $A_1(2) = A_2(1)$ and $A_3(3) = A_3(3)$. None of the rules we defined considers the order of issues in $\mathcal{A}$ when selecting the collective sequences, therefore by construction our rules satisfy neutrality.

The property of independence of irrelevant information states that the collective judgment on each issue depends only on the individual judgments for that issue, and not on the judgments rendered for the other issues in the agenda.

We call two matrixes M_1 and M_2 $j-equal$ when $M_1 \triangledown j = M_2 \triangledown j$. Two profiles π_1 and π_2 over the same N, $\mathcal{A}$ and $\mathcal{R}$ are a_j-*equal* when $\pi_1 \triangledown j = \pi_2 \triangledown j$ for a $a_j \in \mathcal{A}$. We illustrate the property of independence of irrelevant information using the a_1-equal profiles π_1 and π_2 of three agents for agenda $\mathcal{A} = \{a_1, a_2, a_3\}$.

$$\pi_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \qquad \pi_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

It holds $\pi_1 \triangledown a_1 = \pi_2 \triangledown a_1$ and $\pi_1 \triangledown a_2 \neq \pi_2 \triangledown a_2$. Let $f(\pi_1) = A_1$ and $f(\pi_2) = A_2$. If f satisfies the independence of irrelevant information, then $A_1 \triangledown a_1 = A_2 \triangledown a_1$.

There are many ways in which we can lift the function aggregator definition of the independence of irrelevant information property. The first one is by considering a bijective relation between the outputs of the aggregator on two a_j-equal profiles.

Definition 55. *A judgment aggregation rule F satisfies (III-1) when for each a_j-equal profiles and weight matrices, π_1, π_2 and W_1, W_2 correspondingly there exists a bijection $b_f : F(\pi_1, W_1) \mapsto F(\pi_2, W_2)$ such that if $b_f(A) = A'$, then $A \triangledown j = A' \triangledown j$.*

Another way to lift the independence of irrelevant information is by assuring that a collective judgment for a_j, assigned according to some collective sequence for the profile π_1, is also included in some collective sequence for π_2.

Definition 56. *A judgment aggregation rule F satisfies (III-2) when for each j-equal profiles and weight matrices, π_1, π_2 and W_1, W_2 correspondingly, if $A \triangledown j = x$ for all $A \in F(\pi_1)$ then $A' \triangledown j = x$ for all $A' \in F(\pi_2)$.*

When (III-2) is satisfied, (III-1) may not be. Consider $\mathcal{A} = \{p, q, p \wedge q, r\}$, $\mathcal{R} = \emptyset$, $N = \{1, 2, 3\}$, and the following r-equal profiles:

	p	q	$p \wedge q$	r
Agent 1	1	1	1	0
Agent 2	0	1	0	1
Agent 3	1	0	0	1
$F(\pi_1, U)$	0	1	0	1
	1	0	0	1

	p	q	$p \wedge q$	r
Agent 1	0	0	0	0
Agent 2	0	0	0	1
Agent 3	0	0	0	1
$F(\pi_2, U)$	0	0	0	1

In this case (III-2) is satisfied, but (III-1) is not because a bijection cannot exist between sets of different cardinality. If (III-1) is satisfied, then (III-2) is necessarily satisfied. Consider these other r-equal profiles, now for $N = \{1,2,3,4\}$:

	p	q	$p \wedge q$	r
Agent 1	1	0	0	0
Agent 2	1	0	0	0
Agent 3	0	1	0	1
Agent 4	1	1	1	1
$F(\pi_1, U)$	1	0	0	0
	0	1	0	1

	p	q	$p \wedge q$	r
Agent 1	0	0	0	0
Agent 2	0	0	0	0
Agent 3	0	0	0	1
Agent 4	0	0	0	1
$F(\pi_2, U)$	0	0	0	0
	0	0	0	1

In this case, (III-2) is the one that is satisfied. The property (III-1) is trivially satisfied since the condition that $A \triangledown j = x$ for all $A \in F(\pi_1)$ fails. Other lifting of the independence of irrelevant information can be constructed as well. We do not dwell further on the independence of irrelevant information since the construction of our rules is such that we can expect none of them to satisfy the introduced (III-1) and (III-2). As an illustration, consider an example for $\Delta^{d_H,\Sigma}$. Consider the the set $\mathbb{A} = \{(1,0,0),(0,1,0),(0,0,0),(1,1,1)\}$ and the profiles π and π^*:

$$\pi = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \qquad \pi^* = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

We have that $\mathcal{A}_i \triangledown a_2 = A_i^* \triangledown a_3$, for all $i \in \{1,2,3\}$. However, $\Delta^{d_H,\Sigma}(\pi, U) = \{A\} = \{(0,0,0)\}$, while $\Delta^{d_H,\Sigma}(\pi^*, U) = \{A_1^*, A_2^*, A_3^*\} = \{(0,1,0),(1,0,0),(1,1,1)\}$.

4.5 Anonymity

Another property considered among the first in judgment aggregation is that of anonymity. The property of anonymity states that the outcome of the function aggregator does not change regardless of how one permutes the order of the judgment sets in the profile. Since this property does not hinge on the cardinality, input and output type-value of the function aggregator, we can construct the most simple lifting. We begin by defining when a matrix is a permutation of another matrix.

Definition 57. *Let $M_{n\times m}$ and $M'_{n\times m}$ be matrices, let $\mathbf{x} = \langle M \triangleright 1, \ldots, M \triangleright n \rangle$ and $\mathbf{x}' = \langle M' \triangleright 1, \ldots, M' \triangleright n \rangle$. M' is a permutation σ of M, denoted $[M]_\sigma$ if and only if $\mathbf{x}' = [\mathbf{x}]_\sigma$, for a permutation σ.*

Definition 58. *A judgment aggregation rule F from Definition 50 satisfies* anonimity *if for every $\pi \in S_I^n, W \in (\mathbb{R}^+)^{n\times m}$, and every permutation σ, $F(\pi, W) = F([\pi]_\sigma, [W]_\sigma)$.*

The rules based on minimization are anonymous because the row order of the matrix is not considered when selecting the collective output (and $[U]_\sigma = U$ for every σ). In the case of the rules from Chapter 3, whether the rule is anonymous depends on the aggregation function used to construct it.

Proposition 4.5.1. *If $\odot$ is a symmetric aggregation function (see Definition 31), then $\Delta^{d,\odot}$ satisfies anonymity.*

Proof. Assume that $\odot$ is symmetric function. A function $\odot$ is symmetric if and only if $\odot(\mathbf{x}) = \odot([\mathbf{x}]_\sigma)$ for every $\mathbf{x} \in \mathbb{I}^n$ and permutation σ. It follows that $\odot(x_1,\ldots,x_n) = \odot(y_1,\ldots y_n)$ for any two sequences of rational numbers $(x_1,\ldots,x_n)$ and $(y_1,\ldots,y_n)$ such that $\mathbf{y}$ is some permutation σ of $\mathbf{x}$.

Consider a $\pi \in \mathbb{A}^n$ and $W \in (\mathbb{R}^+)^{n\times m}$ and a permutation σ. Let $A \in \mathbb{A}$ and let $s(A,\pi,W)$ denote the score of A with respect to $\pi \triangleright i$ and W calculated as

$$s(i) = \circledast_{j=1}^{m} w_{i,j} \cdot \delta(A(j), \pi_{i,j} \triangleright 1).$$

The scores for $[\pi]_\sigma$ and $[W]_\sigma$ are denoted $s^\sigma(i)$. The sequences $\langle s(1), \ldots, s(n) \rangle$ and $\langle s^\sigma(1),\ldots,s^\sigma(n)\rangle$ are permutations of each other. Therefore, it follows from the symmetry of $\odot$ that

$$\odot(s(1),\ldots,s(n)) = \odot(s^\sigma(1),\ldots,s^\sigma).$$

Consequently, $A \in \Delta^{d,\odot}(\pi,W)$ if and only if $A \in \Delta^{d,\odot}([\pi]_\sigma,[W]_\sigma)$. □

4.6 Adherence to majority

In Chapter 2 we already introduced one relational property, the majority-preservation. A judgment aggregation rule satisfies majority-preservation if a judgment aggregation rule always selects as a collective set the issue-majoritarian set whenever this set exists and is consistent, with respect to $\mathcal{R}$.

There are two possible ways of qualifying a collective judgment set in terms of majority in individual judgments. The first way is to look at the individual judgment sets as atomic information. In this case, a collective judgment set adheres to the majority if more than half of the agents have this judgment set as their individual set. In the literature of distance-based belief merging (Konieczny and Pino-Pérez, 1999, 2002, 2005) such majority-adherence rules are called *majoritarian*. More-precisely, an operator is defined to be majoritarian when there is a number k, such that when k agents have the same belief base, the result of the merging includes that belief base.

The second approach is to consider the judgment set as a divisible and the individual judgments as the atomic information. In this case, a collective judgment set adheres to the majority if, for each judgment in it, there are more than half agents who rendered that judgment. This is how we defined the property of majority-preservation in Section 2.1.1.

The majority-preservation property in binary judgment aggregation, as we define it in Section 2.1.1, corresponds to a well-known property in voting theory, the *Condorcet winner property*. In this section we begin with a discussion on the relations between majority-preservation in binary judgment aggregation and the Condorcet winner property in voting theory.

In the presence of abstentions there are many different notions of majority on a single issue. For each of these notions we can define an issue-majoritarian set and study if a distance-judgment aggregation rule selects it as a collective judgment set. In this section we discuss different majority notions and consider two majority adherence properties, majority-preservation and majoritarianism, for a rule aggregator defined in Definition 50. We study which of the rules from Chapter 3 are adherent to majority.

4.6.1 Condorcet winner property and judgment aggregation

A voting rule satisfies the Condorcet winner property if it selects as a winner the candidate that defeats every other candidate in a pairwise comparison, whenever such candidate exits (Condorcet, 1785; Young and Levenglick, 1978). What is the counterpart of Condorcet winner in judgment aggregation?

In a judgment aggregation context, the Condorcet winner property cannot be directly considered since no preferences between judgments or judgment sets are supplied. However, the translation of a voting problem to a judgment aggregation problem used in (Dietrich and List, 2007a) can be used to "translate" the Condorcet winner property. This translation is for judgment aggregation in which only binary judgments are allowed. For each pair of options a and b we use a proposition p that is true when a is preferred to b and false when b is preferred to a. In this manner only strict preference orders can be translated.

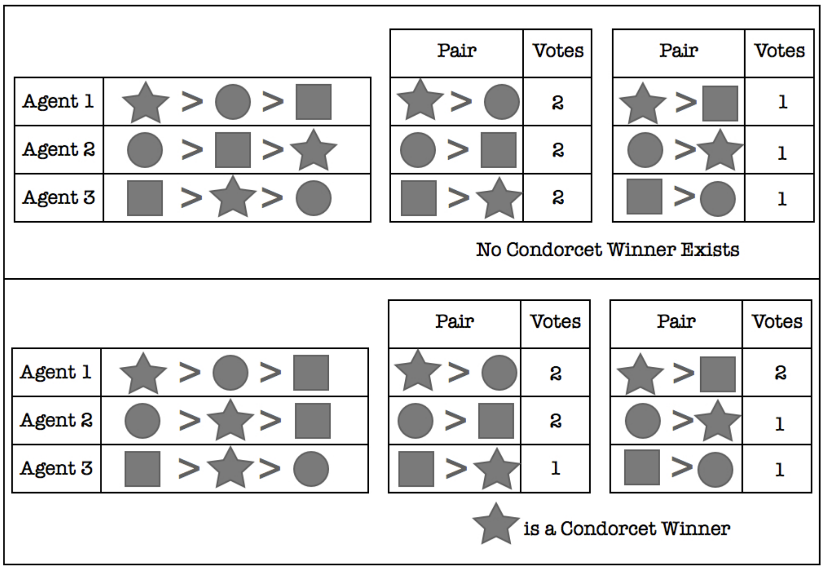
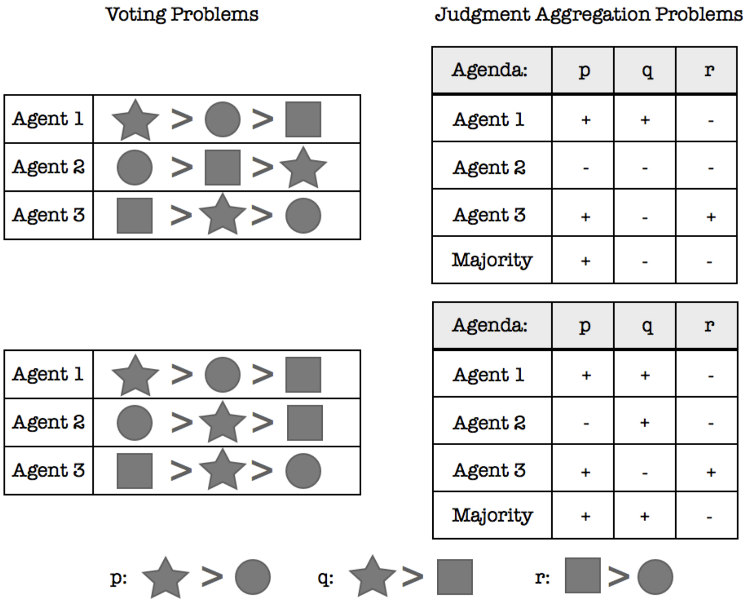


Figure 4.2: Illustration of a Condorcet winner.

Consider as an example three agents that choose from among three options: the star, the square and the circle. In this case we need three propositions in the agenda: p_1 to denote that the star is preferred to the square, p_2 to denote that the star is preferred to the circle and p_3 to denote that the square is preferred to the circle. For instance, the rejection of p_2 denotes that the circle is preferred to the star. In color (on the left-hand side), is the profile of individual preferences. Figure 4.3 illustrates how the preferences from Figure 4.2 are transformed into a judgment aggregation profile using this translation. While for the top profile no Condorcet winner exists, the Condorcet winner of the bottom profile is the star. In gray, on the right-hand side we give the pairwise comparisons. As it can be observed in the figure, the judgment set that contains all the majority-supported judgments corresponds to a preference order in which the top ranked alternative is the Condorcet winner.



Agenda:	p	q	r
Agent 1	+	+	-
Agent 2	-	-	-
Agent 3	+	-	+
Majority	+	-	-

Agenda:	p	q	r
Agent 1	+	+	-
Agent 2	-	+	-
Agent 3	+	-	+
Majority	+	+	-

Figure 4.3: Transforming the voting problem into a judgment aggregation problem. The Condorcet winner corresponds to the proposition majoritarian judgment set.

The majority-preservation property in Definition 6 is the judgment aggregation counterpart, for binary judgment aggregation problems, of the Condorcet winner property, as it was also observed by Nehring et al. (2011).

4.6.2 Majoritarian rules

Let us consider a judgment sequence to be an indivisible whole. A rule aggregator is majoritarian when it necessarily selects, as a collective judgment sequence, the sequence supported by the majority of agents, when such a sequence exists. We only define the majoritarian property for underweighted judgments, namely when $W = U$. When $W \neq U$ then the notion of majoritarian rule is difficult to define if the weights are not agent-associated. To define when a rule is majoritarian for weighted judgments, the weights should be associated with an agent. When weights are associated with issues or with judgments, then these should be transformed into agent associated weights. It is not straightforward how this transformation should be done. Simply summing up the weights on the judgment or on the issue could be one way. Another way is to find the average of the judgment weights for one agent. Yet a third way is to consider the maximum weight assigned to a judgment made by an agent. The minimum of the judgment weights can also be considered a good candidate, etc.

We can consider at least two notions of a sequence being supported by a majority in a profile. The first is the *absolute* majority and the second is the *simple majority*.

A sequence A is supported by an absolute majority, in a profile for n agents, if there are more than half of the agents that selected A than any other sequence in $\mathbb{A}$. We define this formally.

Definition 59. *Let $\pi \in \mathbb{A}^n$. An absolute majority is a partial function $M_a : \mathbb{A}^n \mapsto \mathbb{A}$ defined as*

$$M_a(\pi) = \begin{cases} A & \text{iff } \#\{i|\pi \triangleright i = A\} > \lfloor \frac{n}{2} \rfloor + 1 \\ undefined & otherwise \end{cases}$$

A sequence A is supported by a simple majority, when there are more agents in the profile supporting A when compared to any other sequence $\pi \triangleright i$ in the profile.

Definition 60. *Let $\pi \in \mathbb{A}^n$. A simple majority is a partial function $M_s : \mathbb{A}^n \mapsto \mathbb{A}$ defined as*

$$M_s(\pi) = \begin{cases} A & \text{iff } \#\{i|\pi \triangleright i = A\} > \#\{i|\pi \triangleright i = A'\} \text{ for all } A' \in \mathbb{A} \text{ s.t. } A' \neq A \\ undefined & otherwise \end{cases}$$

If a sequence is supported by an absolute majority, then it is also supported by a simple majority, but the reverse does not hold.

Example 4.6.1. *Consider the profiles π_1, π_2 and π_3 on Figure 4.4.*

$$\pi_1 = \begin{array}{c} \\ E_1,E_2 \\ E_3 \\ E_4 \\ E_5 \end{array} \begin{array}{c} \begin{array}{ccccc} a_1 & a_2 & a_3 & a_4 & a_5 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & \frac{1}{2} & 1 & \frac{1}{2} & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \end{array} \quad \pi_2 = \begin{array}{c} \\ E_1,E_2 \\ E_3,E_4 \\ E_5 \end{array} \begin{array}{c} \begin{array}{ccccc} a_1 & a_2 & a_3 & a_4 & a_5 \end{array} \\ \begin{bmatrix} 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & \frac{1}{2} & 1 & \frac{1}{2} & 1 \end{bmatrix} \end{array} \quad \pi_3 = \begin{array}{c} \\ E_1,E_2,E_3 \\ E_4 \\ E_5 \end{array} \begin{array}{c} \begin{array}{ccccc} a_1 & a_2 & a_3 & a_4 & a_5 \end{array} \\ \begin{bmatrix} 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & \frac{1}{2} & 1 & \frac{1}{2} & 1 \end{bmatrix} \end{array}$$

Figure 4.4: Examples for different notions of majority.

The sequence $\langle 1,0,0,1,1 \rangle$ is supported by a simple majority in π_1, while the absolute majority is undefined for this profile. For the profile π_2 both majorities are undefined. The sequence $\langle 0,0,1,0,1 \rangle$ is supported by both the simple and the absolute majority for π_3.

From Example 4.6.1 we can observe that it does not take many agents for a sequence to be supported by a simple majority, in fact two are enough. The simple majority is a very weak notion when the number of available sequences is comparably larger than the number of agents. Therefore we define the property of majoritarianism using the absolute majority notion.

Intuitively, a judgment aggregation rule is majoritarian if for every profile π it selects as a collective sequence the sequence $M_a(\pi)$ whenever $M_a(\pi)$ is defined. Consider the property of majority-preservation we give in Chapter 2. If a rule R, is majority-preserving, then it is necessarily majoritarian. Therefore we give the formal definition of majoritarianism for the distance-based judgment aggregation rules.

Definition 61. *A judgment aggregation rule $\Delta^{d,\odot}$ is majoritarian when for all $\pi \in \mathbb{A}^n$, if $M_a(\pi)$ exists, then $M_a(\pi) \in \Delta^{d,\odot}$.*

A weaker form of majoritarianism can also be defined by requiring $M_a(\pi) = \Delta^{d,\odot}$ instead of $M_a(\pi) \in \Delta^{d,\odot}$.

Not all aggregation functions $\odot$ give rise to majoritarian rules.

Proposition 4.6.2. *If d is a metric, then $\Delta^{d,\Sigma}$ is majoritarian.*

Proof. Assume, without loss of generality, since $\sum$ is symmetric, that for the first $k = \frac{n}{2}+1$ agents i it is the case that $\pi \triangleright i = A_i = A$, while $A_j \neq A$ for $j \neq i$. We need to show that, for all $A' \in \mathbb{A}$,

$$\sum_{i=k+1}^{n} d(A,A_i) \leqslant (\frac{n}{2}+1)\cdot d(A',A) + \sum_{i=k+1}^{n} d(A',A_i). \tag{4.1}$$

$$\sum_{i=k+1}^{n} d(A,A_i) \leqslant 2d(A',A) + (\frac{n}{2}-1)\cdot d(A',A) + \sum_{i=k+1}^{n} d(A',A_i). \tag{4.2}$$

Since d is a metric, for every i it holds that

$$d(A,A_i) \leqslant d(A',A) + d(A',A_i), \text{ and consequently}$$

$$\sum_{i=k+1}^{n} d(A,A_i) \leqslant \sum_{i=k+1}^{n} d(A',A) + \sum_{i=k+1}^{n} d(A',A_i).$$

The inequality 4.2 can be rewritten as follows:

$$\sum_{i=k+1}^{n} d(A,A_i) \leqslant 2d(A',A) + (\frac{n}{2}-1)\cdot d(A',A) + \sum_{i=k+1}^{n} d(A',A_i). \tag{4.3}$$

Observe that

$$(\frac{n}{2}-1)\cdot d(A',A) = \sum_{i=k+1}^{n} d(A',A) \text{ for } k = \tfrac{n}{2}+1.$$

Since $d(A',A) > 0$, the inequality 4.3 is satisfied. □

Proposition 4.6.3. *If d is a metric, then Δ^{d,Π^*} is majoritarian.*

Proof. When $x,y,z \in \mathbb{R}^+$, if $x \leqslant y+z$, then $x+1 \leqslant (y+1)+(z+1)$ and $x+1 \leqslant (y+1)\cdot(z+1)$. Therefore, the proof of Proposition 4.6.2 can be used to prove this proposition as well.

Assume, without loss of generality since Π^* is symmetric, that for the first $k = \frac{n}{2}+1$ agents i it is the case that $A_i = A$, while $A_j \neq A$ for $j \neq i$. We need to show that, for all $A' \in \mathbb{A}$,

$$\prod_{i=k+1}^{n} (1+d(A,A_i)) \leqslant \cdot(1+d(A',A))^{(\frac{n}{2}+1)} \cdot \prod_{i=k+1}^{n} (1+d(A',A_i)) \tag{4.4}$$

Since d is a metric, for every i it holds that

$d(A,A_i) \leqslant d(A',A) + d(A',A_i)$, but also, since the metrics are positive non-null numbers

$$(1+d(A,A_i)) \leqslant (1+d(A',A)) \cdot (1+d(A',A_i)), \text{and}$$
$$\prod_{i=k+1}^{n}(1+d(A,A_i)) \leqslant \prod_{i=k+1}^{n}(1+d(A',A)) \cdot \prod_{i=k+1}^{n}(1+d(A',A_i)).$$

The inequality 4.4 can be rewritten as follows:

$$\prod_{i=k+1}^{n}(1+d(A,A_i)) \leqslant \cdot(1+d(A',A))^2 \cdot \prod_{i=k+1}^{n}(1+d(A',A)) \cdot \prod_{i=k+1}^{n}(1+d(A',A_i)) \tag{4.5}$$

□

Proposition 4.6.4. $\Delta^{d,max}$ *is not majoritarian.*

Proof. Consider the profile in the proof of Proposition 3.4.4. Two out of three agents select the sequence $\langle 1,1,0\rangle$, however, $\Delta^{d_H,max}(\pi,U) = \{\langle 1,0,0\rangle\}$. □

Corollary 4.6.5. $\Delta^{d,Gmax}$ *is not majoritarian.*

Proof. A consequence of $\Delta^{d,Gmax}(\pi,W) \subseteq \Delta^{d,max}(\pi,W)$. □

The propositions we show are unsurprising since they have been shown to hold for distance-based belief merging operators, see for instance (Konieczny and Pino-Pérez, 2002).

4.6.3 Majority-preservation

It is straightforward to extend the majority-preservation property from unweight to weighted binary judgments. Instead of counting how many agents support a given judgment, we need to consider the sum of the weights of the agent who support the particular judgment. The challenge is in defining the majority-preservation property for ternary, and in principle multi-valued, judgment profiles. The reason for this challenge lies in the many possible ways in which majority on an issue can be defined in the multi-valued case.

Consider for instance the judgments in $\pi\triangledown a_1$ and $\pi\triangledown a_2$ in Figure 4.5. Two judgment majority functions we define in Section 3.5. The first, m_1: is the biased-majority from Definition 44. This majority considers only the number of agents who accept and reject an issue but not those who abstain on an issue. It is defined for every $\pi\triangledown i$ ternary vector of judgments. Using this majority on the issues a_1 and a_2 from Figure 4.5, one accepts both a_1 and a_2. This majority is biased against the judgment $\frac{1}{2}$. Another way to define judgment majority is m_2 from Definition 45: a majority supports a judgment on an issue if there are more agents that select this judgment then any other for that issue. Using this judgment majority on the issues a_1 and a_2 from Figure 4.5, one accepts a_1 but abstains regarding a_2. This majority is not biased against the judgment $\frac{1}{2}$, but it can be considered biased in favor of $\frac{1}{2}$ since both the judgment $\frac{1}{2}$ and 1 for a_2 are supported by an equal number of agents. This majority is also defined for every $\pi\triangledown i$ vector of ternary judgments.

More versions of total judgment majority functions can be constructed. For instance, a function can be defined to combine the biased and unbiased majority based on the number k of agents that abstain on the issue in question:

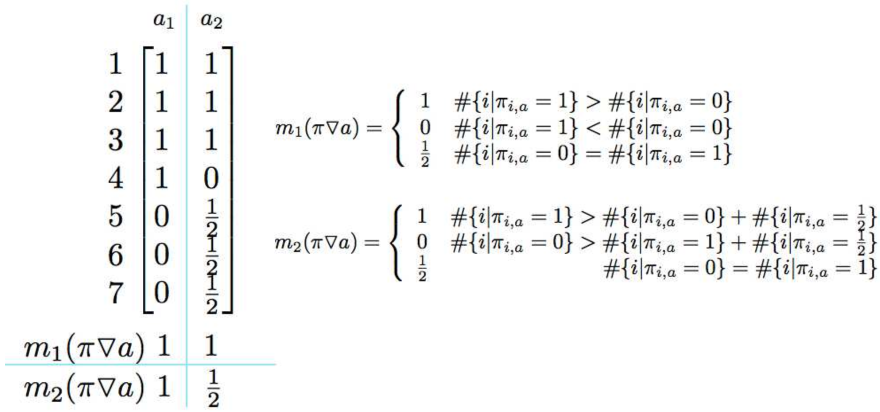


Figure 4.5: Illustration of the different ways to define majority on a single issue.

$$m_3(\pi\nabla a) = \begin{cases} m_1(\pi\nabla a) & \text{iff} \quad \#\{i|\pi_{i,a} = \frac{1}{2}\} \leqslant k \\ m_2(\pi\nabla a) & \text{iff} \quad \#\{i|\pi_{i,a} = \frac{1}{2}\} > k \end{cases}$$

The m_3 can also be considered to be a quota rule (Dietrich and List, 2007b).

We can also define the simple and the absolute majority for judgments on an issue. The intuition behind these two majority functions is the same as in the previous section, where we define them for judgment sequences. A judgment is supported by a simple majority when it is supported by more agents than any other judgment in a pair-wise comparison. A judgment is supported by an absolute majority when there are strictly more than half of the agents supporting it. The un-biased majority can be obtained from the absolute majority when $\frac{1}{2}$ is assigned to all profiles for which the absolute majority is undefined. The simple majority supported judgment is the same as the absolute majority supported judgment, whenever both are defined. All these different notions of majority collapse into one when the judgments are binary.

If a judgment is supported by a simple majority, then it is supported by an absolute majority, and also by an un-biased majority. The judgment majority functions are voting rules applied to choose from the set of options $\{0, \frac{1}{2}, 1\}$. The judgment supported by a simple majority is in fact a Condorcet winner. For these reasons we use, and give the definition of, the weighted simple majority on a judgment.

Definition 62. *Let $\pi \in \mathbb{A}^n$, N a set of agents, $\mathcal{A}$ an agenda, a set of values $T = \{0, \frac{1}{2}, 1\}$ and $x \in \{0, \frac{1}{2}, 1\}$. We define the set $N_x(j) = \{i \mid \pi_{i,j} = x\}$ and the value $V_x(j) = \sum_{i \in N_x(j)} w(i,j)$. The simple judgment majority, on $a_j \in \mathcal{A}$, is a function $m_s : \mathcal{A} \times \mathbb{A}^n \times (\mathbb{R}^+)^{n \times m} \mapsto \{0, \frac{1}{2}, 1\}$ defined as:*

$$m_s(a_j, \pi, W) = \begin{cases} x & \textit{iff} \;\; V_x(j) > V_y(j) \textit{ for any } y \in \{0, \frac{1}{2}, 1\},\ x \neq y \\ \textit{undefined} & \textit{otherwise} \end{cases}$$

For each judgment majority function, we can define a corresponding issue-majoritarian sequence and also a majority-preservation property.

Definition 63. *Let m be a judgment majority function. Given a profile π, for an agenda $\mathcal{A}$ and weight matrix W, the issue majority set $Maj(\pi,W) = \langle m(a,\pi,W) \mid a \in \mathcal{A}\rangle$. If there is $\pi \triangledown j$ such that $m(a_j,\pi,W)$ is undefined, then $Maj(\pi,W)$ does not exist.*

We can further distinguish between strong and weak majority-preservation. A rule is strongly majority-preserving when it selects the set $Maj(\pi,W)$ as a unique collective judgment set for π and W, whenever $Maj(\pi,W)$ is consistent and exists. A rule is weakly majority-preserving if it includes $Maj(\pi,W)$ among the collective judgment sets, whenever $Maj(\pi,W)$ is consistent and exists.

Definition 64 (Weak and strong majority-preservation)**.** *A judgment aggregation rule $F : S_I^n \times (\mathbb{R}^+)^{n\times m} \mapsto S_{II}$ is* strongly majority-preserving *when, if $Maj(\pi,W)$ exists and $Maj(\pi,W) \in S_{II}$, then $F(\pi,W) = Maj(\pi,W)$. The rule F is* weakly majority-preserving *when, if $Maj(\pi,W)$ exists and $Maj(\pi,W) \in S_{II}$, then $Maj(\pi,W) \in F(\pi,W)$.*

As implied from their definitions, the weak majority-preservation is satisfied whenever the strong majority-preservation is satisfied, but the reverse implication does not hold. The rules we defined in Chapter 2, with the exception of $R^{d_H,max}$, satisfy the strong majority-preservation property. In this section we show which of the $\Delta^{d,\odot}$ aggregators satisfy the weak and strong, majority-preservation, with respect to the the simple judgment majority function. From the specific aggregation functions and metrics we considered, we obtain rules that are either strongly majority-preserving or do not satisfy this property at all.

We first prove that $\Delta^{d_H,\Sigma}$ and $\Delta^{d_T,\Sigma}$ are strongly majority-preserving. The proof presented is more detailed than needed, but we construct it in this manner to be able to use it later to build a conjecture regarding what characteristics of $\odot$ and d give rise majority-preserving $\Delta^{d,\odot}$.

Proposition 4.6.6. *If $\odot = \circledast = \sum$ then $\Delta^{d,\odot}$ satisfies the simple strong majority preservation for δ_H and δ_T.*

Proof. We first prove that simple strong majority-preservation holds for single issue agendas and than generalize to arbitrary large agendas.

Let us first consider an agenda with one issue $\mathcal{A} = \{a_1\}$. Let $x,y,z \in \{0,\frac{1}{2},1\}$ such that $x \neq y$, $y \neq z$ and $x \neq z$.

If $\langle x\rangle = \Delta^{\Sigma,d}(\pi,W)$ then (4.6) holds.

$$\sum_{i=1}^{n}\sum j = 1^m(w(i,j)\cdot\delta(x,p_{i,j})) < \sum_{i=1}^{n}\sum j = 1^m(w(i,j)\cdot\delta(y,p_{i,j})) \tag{4.6}$$

Note that (4.6) holds also when z is used instead of y. We can simplify (4.6) as (4.7).

$$\sum_{i=1}^{n}(w(i,1)\cdot\delta(x,p_{i,1})) < \sum_{i=1}^{n}(w(i,1)\cdot\delta(y,p_{i,1})) \tag{4.7}$$

Since $\sum$ is an associative function, we can rewrite (4.7) as (4.8).

$$\sum(\sum_{i\in N_x}(w(i,1)\cdot\delta(x,x)),\sum_{i\in N_y}(w(i,1)\cdot\delta(x,y)),\sum_{i\in N_z}(w(i,1)\cdot\delta(x,z)))<$$
$$\sum(\sum_{i\in N_x}(w(i,1)\cdot\delta(y,x)),\sum_{i\in N_y}(w(i,1)\cdot\delta(y,y)),\sum_{i\in N_z}(w(i,1)\cdot\delta(y,z))) \quad (4.8)$$

We can rewrite (4.8) as (4.9).

$$\sum(0,\delta(x,y)\cdot\overbrace{\sum_{i\in N_y}w(i,1)}^{V_y(a_1)},\delta(x,z)\cdot\overbrace{\sum_{i\in N_z}w(i,1)}^{V_z(a_1)})<\sum(0,\delta(x,y)\cdot\overbrace{\sum_{i\in N_x}w(i,1)}^{V_x(a_1)},\delta(y,z)\cdot\overbrace{\sum_{i\in N_z}w(i,1)}^{V_z(a_1)}) \quad (4.9)$$

If $\delta=\delta_H$, then $\delta(x,y)=\delta(y,z)=\delta(x,z)=1$ and (4.9) becomes (4.10).

$$\sum(0,V_y(a_1),V_z(a_1))<\sum(0,V_x(a_1),V_z(a_1)) \quad (4.10)$$

From (4.10) (and non-decreasing of $\sum$) follows (4.11).

$$V_y(a_1)<V_x(a_1) \quad (4.11)$$

Since (4.10) holds if we swap z and y, (4.12) also follows from (4.10).

$$V_z(a_1)<V_x(a_1) \quad (4.12)$$

If the inequalities (4.11) and (4.12) hold, then $Maj(\pi,W)$ exists and $Maj(\pi,W)=\langle x\rangle$.

Consider $\delta=\delta_T$. We have the following cases:
(case1.1): $x=0$, $y=\frac{1}{2}$ and $z=1$
(case1.2): $x=0$, $y=1$ and $z=\frac{1}{2}$
(case2.1): $x=\frac{1}{2}$, $y=1$ and $z=0$
(case2.2): $x=\frac{1}{2}$, $y=0$ and $z=1$
(case3.1): $x=0$, $y=1$ and $z=\frac{1}{2}$
(case3.2): $x=0$, $y=\frac{1}{2}$ and $z=1$

We apply each of these cases on (4.10), obtaining (4.13), each line corresponding to each case.

$$\begin{aligned}
\frac{1}{2}\cdot V_y(a_1)+V_z(a_1) &< \frac{1}{2}\cdot V_x(a_1)+\frac{1}{2}\cdot V_z(a_1)\\
V_y(a_1)+\frac{1}{2}\cdot V_z(a_1)) &< V_x(a_1)+\frac{1}{2}\cdot V_z(a_1)\\
\frac{1}{2}\cdot V_y(a_1)+\frac{1}{2}\cdot V_z(a_1)) &< \frac{1}{2}\cdot V_x(a_1)+V_z(a_1)\\
\frac{1}{2}\cdot V_y(a_1)+\frac{1}{2}\cdot V_z(a_1)) &< \frac{1}{2}\cdot V_x(a_1)+V_z(a_1)\\
V_y(a_1)+\frac{1}{2}\cdot V_z(a_1)) &< V_x(a_1)+\frac{1}{2}\cdot V_z(a_1)\\
\frac{1}{2}\cdot V_y(a_1)+V_z(a_1)) &< \frac{1}{2}\cdot V_x(a_1)+\frac{1}{2}\cdot V_z(a_1)
\end{aligned} \tag{4.13}$$

We can simplify (4.13) into (4.14).

$$\begin{aligned}
V_y(a_1) &< V_x(a_1)-V_z(a_1)\\
V_y(a_1) &< V_x(a_1)\\
V_y(a_1) &< V_x(a_1)+V_z(a_1)\\
V_y(a_1) &< V_x(a_1)+V_z(a_1)\\
V_y(a_1) &< V_x(a_1)\\
V_y(a_1)+V_z(a_1)) &< V_x(a_1)
\end{aligned} \tag{4.14}$$

In all cases except (case2.1) and (case2.2), we can use the same reasoning as in the case of δ_x to conclude that x is supported by a simple majority in π if defined.

The case (case2.1) gives the same inequality as (case2.2). Since we can swap y and z, for (case2.1) and (case2.2) we obtain that

$$\begin{aligned}
V_y(a_1) &< V_x(a_1)+V_z(a_1)\\
V_z(a_1) &< V_x(a_1)+V_y(a_1)
\end{aligned} \tag{4.15}$$

The (4.15) are possible when $V_z(a_1)=V_y(a_1)$. We can have $V_y(a_1)\geqslant V_x(a_1)$ or $V_y(a_1)<V_x(a_1)$. If $V_y(a_1)\geqslant V_x(a_1)$ then $m_s(a_1,\pi,W)$ is undefined and $M(\pi,W)$ does not exist. If $V_y(a_1)<V_x(a_1)$ then also $V_z(a_1)<V_x(a_1)$ and $m_s(a_1,\pi,W)=x$. Consequently $Maj(\pi,W)=\langle x\rangle$.

Now let us assume that $\mathcal{A}$ has m elements. Assume that $M(\pi,W)$ exists and $M(\pi,W)=A^*$, $A^*\in\mathbb{A}$. From the proofs for $|\mathcal{A}|=1$ we obtain that, for any $A\in\mathbb{A}$, $A\neq A^*$, and every $j\in\{1,\ldots,m\}$, (4.16) holds.

$$\sum_{i=1}^{n} w(i,j)\cdot\delta(A^*(j),p_{i,j})\leqslant\sum_{i=1}^{n} w(i,j)\cdot\delta(A(j),p_{i,j}) \tag{4.16}$$

From (4.16) and the non-decreasingness of $\sum$, we obtain (4.17).

$$\sum j = 1^m \sum_{i=1}^{n} w(i,j)\cdot\delta(A^*(j),p_{i,j}) \leqslant \sum j = 1^m \sum_{i=1}^{n} w(i,j)\cdot\delta(A(j),p_{i,j}) \qquad (4.17)$$

Inequality (4.17) is equivalent to (4.18).

$$\sum_{i=1}^{n}\sum j = 1^m w(i,j)\cdot\delta(A^*(j),p_{i,j}) \leqslant \sum_{i=1}^{n}\sum j = 1^m w(i,j)\cdot\delta(A(j),p_{i,j}) \qquad (4.18)$$

From (4.18) we can conclude that if $M(\pi,W)\in\mathbb{A}$, then $\Delta^{d,\Sigma} = M(\pi,W)$, for $d = d_H$ and $d = d_T$. □

We show that the rest of the examples of distance-based rules we considered do no satisfy the weak, and with that neither the strong, majority-preservation property. For the non majoritarian rules this result is implied.

Proposition 4.6.7. *$\Delta^{d,max}$ and $\Delta^{d,Gmax}$ are not (weakly) majority-preserving.*

Proof. It is sufficient to observe that if a sequence in the profile is supported by an absolute majority of agents then each judgment in that sequence is suported by a simple majority. Therefore, the counter-examples that show that $\Delta^{d,max}$ and $\Delta^{d,Gmax}$ are not majoritarian, are also counter-examples that show that these rules are not majority-preserving. □

More interesting is the case of $\odot = \Pi^*$ and $\Delta^{d_D,\Sigma}$. These rules are majoritarian, but not majority-preserving.

Proposition 4.6.8. *Δ^{d_H,Π^*} and Δ^{d_T,Π^*} are not (weakly) majority-preserving.*

Proof. One binary counter-example suffices for both d_H and d_T. Let $\mathcal{A} = \{a_1,a_2,a_3,a_4\}$ and

$$\mathbb{A} = \left\{ \begin{matrix} \langle 1,1,1,1\rangle, \langle 0,0,0,0\rangle, \langle 1,1,0,0\rangle, \\ \langle 0,0,1,1\rangle,\ \langle 1,0,1,0\rangle \end{matrix} \right\}.$$

Consider the profile

$$\pi = \begin{bmatrix} 1\,1\,1\,1 \\ 1\,1\,1\,1 \\ 0\,0\,0\,0 \\ 0\,0\,0\,0 \\ 1\,1\,0\,0 \\ 0\,0\,1\,1 \\ 1\,0\,1\,0 \end{bmatrix}.$$

We have that $Maj(\pi,U) = \{\langle 1,0,1,0\rangle\}$. However there exists an $A\in\mathbb{A}$ such that

$$\prod_{i=1}^{n}(1+d_H(A,\pi\rhd 1),\dots 1+d_H(A,\pi\rhd n)) < \prod_{i=1}^{n}(1+d_H(A^m,\pi\rhd 1),\dots 1+d_H(A^m,\pi\rhd n)).$$

Namely

$$\prod_{i=1}^{7}(1+d_H(A^m,\pi\triangleright 1),\dots 1+d_H(A^m,\pi\triangleright 7)) = 3\cdot 3\cdot 3\cdot 3\cdot 3\cdot 3\cdot 1 = 729, \text{ but}$$

$$\prod_{i=1}^{7}(1+d_H(\pi\triangleright 1,\pi\triangleright 1),\dots 1+d_H(\pi\triangleright 1,\pi\triangleright 7)) = 1\cdot 1\cdot 1\cdot 5\cdot 3\cdot 3\cdot 3 = 675, \text{ and}$$

$$\prod_{i=1}^{7}(1+d_H(\pi\triangleright 3,\pi\triangleright 1),\dots 1+d_H(\pi\triangleright 3,\pi\triangleright 7)) = 675.$$

□

Proposition 4.6.9. $\Delta^{d_D,\Sigma}$ *is not (weakly) majority-preserving.*

Proof. We show that there is a profile π such that $Maj(\pi,W) \notin \Delta^{d_D,\Sigma}(\pi,W) \neq maj(\pi)$ but $Maj(\pi,W) \in \mathbb{A}$. Let $\mathcal{A} = \{a, a \to (b \wedge c), b, c\}$. The set $\mathbb{A}$ for $\mathcal{A}$ is

$$\mathbb{A} = \left\{ \begin{array}{l} \langle 0,1,0,0\rangle, \langle 0,1,0,1\rangle, \langle 0,1,1,0\rangle, \langle 1,0,1,0\rangle, \\ \langle 0,1,1,1\rangle, \langle 1,0,0,0\rangle, \langle 1,0,0,1\rangle, \langle 1,1,1,1\rangle \end{array} \right\}$$

Consider the following profile, a_1 denotes the expression $a \to (b \wedge c)$:

$$\pi = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array}\!\!\begin{array}{c} a\ a_1\ b\ c \\ \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \end{array}.$$

The $Maj(\pi,U) = \{\langle 0,1,1,1\rangle\}$ is an element of $\mathbb{A}$, however it is not among the outputs of $\Delta^{d_D,\Sigma}(\pi,U)$, as shown on Table 4.1. □

$A \in \mathbb{A}$	$d_D(A,A_1)$	$d_D(A,A_2)$	$d_D(A,A_3)$	$\sum$
$\langle 0,1,0,0\rangle$	0	1	1	2
$\langle 0,1,0,1\rangle$	1	1	1	3
$\langle 0,1,1,0\rangle$	1	1	1	3
$\langle 0,1,1,1\rangle$	1	0	1	2
$\langle 1,0,0,0\rangle$	1	1	1	3
$\langle 1,0,0,1\rangle$	1	1	1	3
$\langle 1,0,1,0\rangle$	1	1	1	3
$\langle 1,1,1,1\rangle$	1	1	0	2

Table 4.1: The sum of Hamming metrics from an element in $\mathbb{A}$ to each of the judgment sequences. $\Delta^{d_D,\Sigma}(\pi,U) = \{\langle 0,1,0,0\rangle, \langle 0,1,1,1\rangle, \langle 1,1,1,1\rangle\}$.

The simple judgment majority function is not defined for every column $\pi\triangledown j$ in a profile. How can we define a total judgment majority function and not violate the majority-preservation of the distance-based rules that satisfy it? To form this question differently, which default judgment should be assigned when the majority is undefined. Since a distance-based merging rule selects the judgment that minimizes distances, the default judgment should be the one in the set of valuations T which is at a minimal distance from each other judgment in T. In the case of δ_T the default judgment should be $\frac{1}{2}$, while in the case of d_H, both 0 and 1 together should be the default judgment.

Example 4.6.10. *Consider for example*[2] *the single-issue profile* π *and weight matrix* W.

$$\pi = \begin{bmatrix} 1 \\ \frac{1}{2} \\ 0 \end{bmatrix} \qquad W = \begin{bmatrix} 5 \\ 1 \\ 5 \end{bmatrix}$$

The simple majority is undefined for this π. *We obtain* $\Delta^{d_T,\Sigma}(\pi,W) = \langle \frac{1}{2} \rangle$, *while* $\Delta^{d_H,\Sigma}(\pi,W) = \{\langle 1 \rangle, \langle 0 \rangle\}$.

We can define the median of T with respect to a distance δ.

Definition 65. *Let* T *be a set of values and let* $\delta : T \times T \mapsto \mathbb{R}^+$ *be a metric. The median of* T *with respect to* δ *is the set* $med_\delta(T) = \{x \mid x \in T$ *and for all* $y,z \in T, \delta(x,z) \leqslant \delta(y,z)\}$.

We can now define a total simple judgment majority mt_s function as:

$$mt_s(a_j,\pi,W) = \begin{cases} x & \text{iff } V_x(j) > V_y(j) \text{ for any } x,y \in \{0,\frac{1}{2},1\},\ x \neq y \\ med_\delta(T) & \text{otherwise} \end{cases}\Bigg)$$

An inevitable question to ask at this point is what are the properties of $\odot$, $\circledast$ and δ that give rise to a majority-preserving $\Delta^{\odot,d}$. It can be conjectured that all majority-preserving distance-based rules $\Delta^{d',\odot}$ are such that $\Delta^{d',\odot} = \Delta^{d,\Sigma}$, for some product metric d constructed using $\circledast = \sum$.

Let us start with an associative aggregation function $\odot$ that satisfies minimality. We obtain (4.19) by re-writing (4.8) using $\odot$ instead of $\sum$.

$$\begin{aligned} &\odot(\odot_{i\in N_x}(w(i,1)\cdot\delta(x,x)),\odot_{i\in N_y}(w(i,1)\cdot\delta(x,y)),\odot_{i\in N_z}(w(i,1)\cdot\delta(x,z))) < \\ &\quad \odot(\odot_{i\in N_x}(w(i,1)\cdot\delta(y,x)),\odot_{i\in N_y}(w(i,1)\cdot\delta(y,y)),\odot_{i\in N_z}(w(i,1)\cdot\delta(y,z))) \end{aligned} \tag{4.19}$$

Inequality (4.19) is equivalent to (4.20).

$$\begin{aligned} &\odot(\odot_{i\in N_y}(w(i,1)\cdot\delta(x,y)),\odot_{i\in N_z}(w(i,1)\cdot\delta(x,z))) < \\ &\quad \odot(\odot_{i\in N_x}(w(i,1)\cdot\delta(x,y)),\odot_{i\in N_z}(w(i,1)\cdot\delta(y,z))) \end{aligned} \tag{4.20}$$

To derive any relation between (4.20) and the simple majority, there must be a relation between $\odot$ and $\sum$. More precisely, there must be an order-preserving mapping between $\odot(\mathbf{x})$ and $\sum(\mathbf{x})$. If such an order-preserving map exists, then $\odot(\mathbf{x}) < \odot(\mathbf{y})$ if and only if $\sum(\mathbf{x}) < \sum(\mathbf{y})$. But in this case, $\Delta^{\odot,d}$ selects the same judgment, on an single-issue agenda,

[2] I thank an anonymous reviewer for providing me with this example.

as $\Delta^{\sum,d}$. As an example, consider the distance $m_P(\mathbf{x},\mathbf{x}') = \Pi_{i=1}^{m} 2^{\delta_H(x_1,x_i')}$ we defined in Section 3.4. This function has $\circledast = \Pi$. Selecting $\odot = \Pi$ we do obtain a majority-preserving rule $\Delta^{m_P,\Pi} = \Delta^{d_H,\sum}(\pi,W)$.

To do the extension from a single-agenda to an arbitrarily large agenda, namely to go from (4.17) to (4.18), we need that $\circledast$ commutes with $\odot$. The only functions $\circledast$ that commute with $\sum$ are such that $\circledast(\mathbf{x})$ is a linear transform of $\sum(\mathbf{x})$. Such a function, for instance, is the arithmetic mean *AM*. In principle, we can always take $\circledast = \odot$ to ensure that $\circledast$ and $\odot$ commute.

The conjecture can be proved by showing the following: there is no majority-preserving $\Delta^{d,\odot}$ for not commuting $\odot$ and $\circledast$.

Although the search for majority-preserving rules is discouraging, the observation made here can be used to construct other distance-based rules. Instead of aggregating row by row, we can design a distance-based rule to aggregate column by column. This is what we accomplish when we swap the order between $\odot$ and $\circledast$, when they commute. The judgment majority rules use the $\sum$ to aggregate the judgments in the column. For instance, in the presence of weights we can use $\circledast = max$ to obtain the judgment on an issue that is associated with the highest weight. Using $\odot = sum$ the distance-based rule aggregator will return the sequence with the highest weighted judgments, whenever such sequence exists and is consistent.

4.7 Unanimity adherence

Unanimity is one of the most natural relational properties in social choice stating that if all agents submit the same individual information to be aggregated, then the aggregate is precisely that information. As in the case of majority-adherence, the judgment sequence can be seen as a whole and agents are unanimous when every agent selects the same judgment sequence. For instance, a sequence-unanimous profile in this sense is:

$$\pi = \begin{bmatrix} 1\,0\,0\,1 \\ 1\,0\,0\,1 \\ 1\,0\,0\,1 \end{bmatrix}.$$

Unanimity is a property satisfied by a function aggregators, defined in *e.g.*, (List and Puppe, 2009), when for every profile π sequence-unanimous on A, $f(\pi) = A$.

The judgment sequence can also be seen as a partitionable collection of judgments and unanimity can be considered in the case of profiles in which the agents are unanimous in their judgments on a given issue. For instance, a judgment-unanimous profile in this sense is:

$$\pi = \begin{bmatrix} 1\,1\,0\,1 \\ 0\,0\,0\,1 \\ 1\,0\,0\,1 \end{bmatrix}.$$

The judgment-unanimous profiles are considered by a property called *unanimity principle* (Dietrich and List, 2008b). The unanimity principle considers whether the unanimously selected judgment is included in the collective judgment sets. For function aggregators, the unanimity principle is defined in (Dietrich and List, 2008b) in the following way. For every profile $\langle A_1,\ldots,A_n\rangle$ in the domain of the aggregation function f and all $\varphi \in \mathcal{A}$, if $\varphi \in A_i$ for all individuals i, then $\varphi \in f(A_1,\ldots,A_n)$.

The unanimity principle is a stronger property than unanimity, since all functions that satisfy the unanimity-principle are also unanimous. We lift both the unanimity and unanimity principle and study when they are satisfied by the rules we introduced.

4.7.1 Unanimity

Unanimity is a relatively weak property in voting theory and preference aggregation, satisfied by virtually all rules. In judgment aggregation, this property is still weak enough to be satisfied by all the rules we introduce.

Definition 66. *A judgment aggregation rule F (Definition 50) is unanimous when for every $W \in (\mathcal{R}^+)^{n\times m}$ if π is such that $\pi \triangleright 1 = \cdots = \pi \triangleright n = A$, then $F(\pi, W) = \{A\}$.*

A profile such that $\pi \triangleright 1 = \cdots = \pi \triangleright n = A$ is majority-consistent, therefore all the majority-preserving rules defined in Chapter 2 are unanimous as well. The rule $R^{d_H,max}$ is a special case of the rule aggregator $\Delta^{d_H,max}$. This aggregator is unanimous as a corollary of the following proposition.

Proposition 4.7.1. *If*

1. *$\circledast$ satisfies minimality (Definition 32);*
2. *$inf\ \odot = \odot(\mathbf{x}) = k$ if and only if $\mathbf{x} = \mathbf{0}$;*

then $\Delta^{d,\odot}$ is unanimous.

Proof. Recall that if $\odot$ satisfies minimality, then it has a unique minimum $k = 0$ for $x_1 = x_2 = \cdots = x_n = 0$. When $A_1 = A_2 = \cdots = A_n = A$ and $\circledast$ satisfies minimality $d(A, A_i) = 0$. If $\odot$ satisfies minimality, then $\odot(d(A, A_1), \ldots, d(A, A_n)) = 0$. Every other judgment sequence will have a score higher than 0. The aggregation function Π^* has a unique minimum in $k = 1$, but since the only values for which $\Pi^*(\mathbf{x}) = 1$ when $\mathbf{x} \in (\mathbb{R}^+)^n$ are $x_1 = \cdots x_2 = 0$, the rule Δ^{d,Π^*} is unanimous as well. □

4.7.2 Unanimity principle

We lift the unanimity principle of (Dietrich and List, 2008b) to two properties which we call the *weak* and the *strong unanimity principle*.

Definition 67. *Let W be some weight matrix. A judgment aggregation rule F satisfies:*

- weak unanimity *(WU) when, for every π such that $\pi_{1,j} = \cdots = \pi_{n,j} = x$ for a $a_j \in \mathcal{A}$, there exists a $A \in F(\pi, W)$ such that $A(j) = x$;*
- strong unanimity *(SU) when, for every π such that $\pi_{1,j} = \cdots = \pi_{n,j} = x$ for a $a_j \in \mathcal{A}$, for all $A \in F(\pi, W)$, $A(j) = x$.*

If (SU) is satisfied by F, then so is (WU).

We show which of the rules based on minimization from Chapter 2 and the distance based rules from Chapter 3 satisfy (WU) and (SU). We consider the rules: R_{MSA} (Definition 11),

R_{MCSA} (Definition 12), R_{MWA} (Definition 15), R_{RA} (Definition 16), R_Y (Definition 17), R_{RY} (Definition 19) and R_{MNAC} (Definition 23). The rule $R^{d_H,max}$ (defined in Section 2.2.4) we consider as part of the distance based rules. We do not analyze the rules R_{IY} (Definition 18) and R_{MR} (Definition 21) since these rules are very weak. Namely these rules select a very large number of collective judgment sets and are more (with respect to their theories) general than most rules.

Proposition 4.7.2. *R_{MSA} satisfies weak unanimity but not strong unanimity.*

Proof. Let P be a profile on an agenda $\overline{\mathcal{A}}$, and $\varphi \in \overline{\mathcal{A}}$ on which all agents give the same judgment x. There always exists a maximal consistent sub-agenda, with respect to set inclusion, that contains φ. Consequently there exists a judgment set in $R_{MSA}(P)$ that contains x.

As a counter-example for $R_{MSA}(P)$ satisfying strong unanimity, consider the profile P of the proof of Proposition 2.3.5. R_{MSA} does not satisfy strong unanimity, because $a \notin T_{R_{MSA}}(P)$.

□

Proposition 4.7.3. *R_{MCSA} does not satisfy weak (or strong) unanimity.*

Proof. Consider again the profile P of the proof of Proposition 2.3.5. The only maxcard consistent sub-agenda of P contains $\neg a$ (and does not contain a). Consequently R_{MCSA} does not even satisfy weak unanimity.

□

Proposition 4.7.4. *R_{MWA} does not satisfy weak (or strong) unanimity.*

Proof. See again the counterexample that can be found in (Pigozzi et al., 2009), which we presented in the proof of Proposition 2.3.8.

□

Proposition 4.7.5. *R_{RA} satisfies strong (and weak) unanimity.*

Proof. Let P be a profile and $Y_P \subseteq \overline{\mathcal{A}}$ be the subset of the agenda consisting of all elements on which there is unanimity among the agents. Because individual judgment sets are consistent, the conjunction of all elements of Y is consistent. Now, when computing $R_{RA}(P)$, the elements of Y are considered first, and whatever the order in which they are considered, they are included in δ because no inconsistency arises. Therefore, for all $\alpha \in Y_P$ and all $\hat{A} \in R_{RA}(P)$, we have $\alpha \in \hat{A}$. □

Proposition 4.7.6. *R_Y satisfies strong (and weak) unanimity.*

Proof. Observe that if α is unanimously accepted by all agents in the set N, it is consequently unanimously selected by all consistent subsets of N. □

Proposition 4.7.7. *R_{RY} satisfies strong (and weak) unanimity.*

Proof. Similar as the proof of Propositions 4.7.6. If α is in all judgment sets for a profile P, then α is in all judgment sets for any super-profile of P constructed by repeating the judgment sets from P. □

Proposition 4.7.8. *R_{MNAC} does not satisfy weak (nor strong) unanimity.*

Proof. Again consider the agenda and profile P in the proof of Proposition 2.3.8. Since $R_{MNAC}(P) = R_{MWA}(P)$, this example of a profile is a counter-example for R_{MNAC} satisfying weak unanimity as well. □

Not all $\Delta^{d,\odot}$ satisfy unanimity. We consider the d given in Section 3.3.2 and the $\odot$ given in Section 3.3.1.

Proposition 4.7.9. *$\Delta^{d_D,\odot}$ satisfies the strong unanimity principle.*

Proof. Recall that the drastic metric is defined as $d(A,A') = 0$ if $A ='$ and $d(A,A') = 1$ otherwise. As a consequence of this definition and the non-decreasing of $\odot$, $A \in \Delta^{d_D,\odot}(\pi,W)$, if and only if there exists an i such that $A = \pi_i$, *i.e.*, A is necessarily in the profile.

□

Proposition 4.7.10. *$\Delta^{d,\Sigma}$ does not satisfy the weak unanimity principle for $d \in \{d_T, d_H\}$.*

Proof. It is sufficient to give a counter-example for $W = U$. We consider the same example as in the proof of Proposition 3.4.5. As it can be observed in Table 3.8, although the agents are unanimous on $A_i \nabla a_{13} = 1$, the only value for a_{13} is $A_4 \nabla a_{13} = 0$. □

Proposition 4.7.11. *$\Delta^{d,max}$ does not satisfy the weak unanimity principle for $d \in \{d_T, d_H\}$.*

Proof. As we did in the proof of Theorem 4.7.10, here also it suffices to give the counter-example for $W = U$. Consider the agenda $\mathcal{A} = \{a_1, a_2, a_3, a_4, a_5, a_6, a_7\}$. The set of consistent judgment sets $\mathbb{A}$ is given in Table 4.2, first column. Let the profile consist of the judgment sets:
$\hat{A}_1 = \{a_1, \neg a_2, \neg a_3, a_4, \neg a_5, \neg a_6, \neg a_7\}$;
$\hat{A}_2 = \{\neg a_1, a_2, \neg a_3, \neg a_4, a_5, \neg a_6, \neg a_7\}$;
$\hat{A}_3 = \{\neg a_1, \neg a_2, a_3, \neg a_4, \neg a_5, a_6, \neg a_7\}$.

It is the case that $\neg a_7 \in \hat{A}_1$, $\neg a_7 \in \hat{A}_2$ and $\neg a_7 \in \hat{A}_3$. However, as it can be observed from Table 4.2, the rule $\Delta^{d_H,max}$ selects a unique judgment set that does not contain $\neg a_7$.

$\hat{A} \in \hat{\mathbb{A}}$	$d_H(\hat{A},\hat{A}_1)$	$d_H(\hat{A},\hat{A}_2)$	$d_H(\hat{A},\hat{A}_3)$	max
$\{a_1, \neg a_2, \neg a_3, a_4, \neg a_5, \neg a_6, \neg a_7\}$	0	4	4	4
$\{\neg a_1, a_2, \neg a_3, \neg a_4, a_5, \neg a_6, \neg a_7\}$	4	0	4	4
$\{\neg a_1, \neg a_2, a_3, \neg a_4, \neg a_5, a_6, \neg a_7\}$	4	4	0	4
$\{\neg a_1, \neg a_2, \neg a_3, \neg a_4, \neg a_5, \neg a_6, a_7\}$	3	3	3	3

Table 4.2: The *max* of Hamming metrics from an element in the set $\hat{\mathbb{A}}$ to each of the agent's judgment sets. The judgment set chosen by the rule $\Delta^{d_H,max}$ does not contain the unanimously selected $\neg a_7$.

□

Proposition 4.7.12. *Δ^{d,Π^*} does not satisfy the weak unanimity principle for $d \in \{d_T, d_H\}$.*

Proof. We give a counter-example for Δ^{d_H,Π^*} when $W = U$. Consider an agenda $\mathcal{A}$. The set of all consistent judgment sequences, for $\mathcal{A}$, is given in the first column (from the left) in Table 4.3.

$A \in \mathbb{A}$
$A_1 = (1,0,0,0,0,0,1,0,0,0,0,0,1,0,0,0,0,0,1,0,0,0,0,0,1)$
$A_2 = (0,1,0,0,0,0,0,1,0,0,0,0,0,1,0,0,0,0,0,1,0,0,0,0,1)$
$A_3 = (0,0,1,0,0,0,0,0,1,0,0,0,0,0,1,0,0,0,0,0,1,0,0,0,1)$
$A_4 = (0,0,0,1,0,0,0,0,0,1,0,0,0,0,0,1,0,0,0,0,0,1,0,0,1)$
$A_5 = (0,0,0,0,1,0,0,0,0,0,1,0,0,0,0,0,1,0,0,0,0,0,1,0,1)$
$A_6 = (0,0,0,0,0,1,0,0,0,0,0,1,0,0,0,0,0,1,0,0,0,0,0,1,1)$
$A_7 = (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0)$

Table 4.3: The set of consistent judgment sentences for $\mathcal{A}$.

Consider the profile π where $\pi \triangleright i = A_i$ for $i \in [1,\ldots,6]$. All the judgment sequences are unanimous on the last issue, assigning it a judgment 1. However, as it can be observed from Table 4.4, the judgment sequence selected by Δ^{d_H,Π^*} is A_7 which assigns a judgment 0 to the last sequence.

$A \in \mathbb{A}$	$d_H(A,A_1)$	$d_H(A,A_2)$	$d_H(A,A_3)$	$d_H(A,A_4)$	$d_H(A,A_5)$	$d_H(A,A_6)$	Π^*
A_1	0	8	8	8	8	8	59 049
A_2	8	0	8	8	8	8	59 049
A_3	8	8	0	8	8	8	59 049
A_4	8	8	8	0	8	8	59 049
A_5	8	8	8	8	0	8	59 049
A_6	8	8	8	8	8	0	59 049
A_7	5	5	5	5	5	5	46 656

Table 4.4: The Π^* of Hamming metrics from an element in the set $\mathbb{A}$ to each of the agent's judgment sequences. The judgment sequence chosen by the rule Δ^{d_H,Π^*} is A_7 for which $A_7 \triangledown a_{25} = 1$ although in the profile $A_i \triangledown a_{25} = 0$ for all i.

□

We can observe that in general, whether the unanimity-principle is satisfied by a distance-based rule does not depend on the properties of d and $\odot$, but rather on the ratio between the cardinality of the agenda on one side, and the number of agents and associated weights for each judgment, on the other side. Let x be the judgment supported by all agents for a in a profile π. Unanimity-preservation will be satisfied always when the minimal value of $\odot_{i=1}^{n} w(i,j)\delta(y,\pi)$ for $y \neq x$, is larger than the minimal value of $\odot_{i=1}^{n} \circledast_{a_j \neq a}^{n} w(i,j)\delta(y,\pi)$. These conditions for unanimity-preserving are very strict, therefore they are not included as a proposition.

4.8 Monotonicity properties

In voting theory monotonicity is a standard property considered for voting rules. When a voting rule is monotonic, an improvement in the ranking of the winning alternative, *ceteris paribus*, does not diminish that alternative's likelihood of being a winner. When the purpose of aggregation is to select an alternative that is representative of the individual input, then it is desirable that additional support for an input should not make that input less likely to be the aggregate (Nurmi, 2004).

In judgment aggregation monotonicity has also been considered as a desirable relational property. There are three versions of monotonicity defined for function aggregators: *monotonicity on an agenda issue* as a property imposed on an aggregation function (List and Puppe, 2009), *monotonicity* as a property imposed on a subset of the agenda (to address manipulability issues) (Dietrich and List, 2005), and *monotonicity on a judgment set* by (Dietrich and List, 2008a). The first property is the strongest, subsuming the other two.

The monotonicity property defined for function aggregators in (List and Puppe, 2009) can be lifted to monotonicity for rule aggregators in the following way.

Definition 68 (Monotonicity)**.** *A profile P' is called an i-variant of profile P when for all $i \neq j$, $i, j \in [1, m]$, $\hat{A}_j \in P$ if and only if $\hat{A}_j \in P'$. A judgment aggregation rule F is monotonic when, for every P, $P' \in S_I^n$ such that $P = (\hat{A}_1, \ldots, \hat{A}_i, \ldots, \hat{A}_n)$ and $P' = (\hat{A}_1, \ldots, \hat{A}_i', \ldots, \hat{A}_n)$ its i-variant, and a $W \in (\mathbb{R}^+)^{n\times m}$ if there is an $a \in \mathcal{A}$ such that*

- $a \notin \hat{A}_i$;
- $a \in \hat{A}_i'$;
- $a \in T_F(P)$;

then $a \in T_F(P')$ and $F(P', W) = F(P, W)$.

The monotonicity property defined as above is a very strong property. There are no constraints imposed on $\hat{A}_i'$ with respect to $\hat{A}_i$, therefore it can happen that $\hat{A}_i' \cap \hat{A}_i = \emptyset$ and for rules that do not satisfy independence the collective judgment set can be affected on more issues than just φ. We can define a weaker monotonicity property and we consider whether our rules satisfy it. The intuition behind our new property is closer to the intuition behind the monotonicity property as studied in voting theory. Namely, the *ceteris paribus* improvement in the support for a judgment that is already included in all collective judgment sets, should not diminish that judgment's likelihood of being in all collective judgment sets. This is the property of *insensitivity to reinforcement of collective judgments*.

First we define when a profile is an α-improvement of another profile for the case of $T \subseteq \{0, \frac{1}{2}, 1\}$.

Definition 69. *Given two profiles $\pi, \pi' \in \mathbb{A}^n$, $\alpha \in \{0, 1\}$ and $a_j \in \mathcal{A}$, the profile π' is called an* α-improvement *of π when*

- *for all $k \neq i$ and for all $r \neq j$ $\pi_{k,r} = \pi'_{k,r}$;*
- $\pi_{i,j} \neq \alpha$;

- $\pi'_{i,j} = \alpha$.

For binary judgments, the condition $\pi_{i,j} \neq \alpha$ implies that $\pi_{i,j} = \neg\alpha$. We do not consider reinforcements for the judgment $\frac{1}{2}$.

Consider for example the agenda $\mathcal{A} = \{a_1, a_2, a_3\}$, $\mathcal{R} = \{(a_1 \wedge a_2) \leftrightarrow a_3\}$, $\models_{3L}$ (Łukasiewicz logic) and $N = \{1,2,3\}$. The profiles π' and π'' are correspondingly a 0-improvement and 1-improvement of π for a_1.

$$\pi = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 1 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \qquad \pi' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \qquad \pi'' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \qquad \pi''' = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 1 \\ 1 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

The profile π''' is not a improvement of π since it is not in $\mathbb{A}^3$.

As with the other relational properties, many versions of insensitivity to reinforcement property can be defined. We construct two versions.

Definition 70 (IR-s)**.** *Let α be a judgment for $a \in \mathcal{A}$. F satisfies* strict insensitivity to reinforcement of collective judgements I *if for all α-improvement profiles $\pi' \in S_I^n$ of $\pi \in S_I^n$ if [for all $A \in F(\pi, W)$, $A(a) = \alpha$] then $F(\pi, W) = F(\pi', W)$.*

Definition 71 (IR)**.** *Let α be a judgment for $a \in \mathcal{A}$. F satisfies* insensitivity to reinforcement of collective judgements I *if for all α-improvement profiles $\pi' \in S_I^n$ of $\pi \in S_I^n$ if [for all $A \in F(\pi, W)$, $A(a) = \alpha$] then [for all $A' \in F(\pi, W)$, $A'(a) = \alpha$] .*

If a rule F is monotonic, then it satisfies (IR-s). If F satisfies (IR-s), then it satisfies (IR).

We show which of the rules based on minimization from Chapter 2 satisfy (IR-s). We do not consider R_{IY} and R_{MR}. The rules from Chapter 3 we analyze with respect to (IR).

Proposition 4.8.1. *R_{MSA} and R_{MCSA} satisfy (IR-s).*

Proof. We consider R_{MSA}. Assume that $\alpha \in T_{R_{MSA}}(P)$ and P' an $\alpha - reinforcement$ of P.

Let $Y \subseteq \overline{\mathcal{A}}$ be a maximal agenda for which $P^{\downarrow Y}$ is majority-consistent. Because $\alpha \in T_{R_{MSA}}(P)$, we must have $\alpha \in Y$. $P'^{\downarrow Y}$ must be majority-consistent as well due to the conditions of Definition 69. Moreover

$$M(P'^{\downarrow Y}) = M(P^{\downarrow Y}). \tag{4.21}$$

From (4.21) it is infered that

all maximal majority-consistent sub-agendas for P' contain some maximal majority-consistent sub-agenda forP'. (4.22)

Let $Y \subseteq \overline{\mathcal{A}}$ be a maximal agenda for which $P'^{\downarrow Y}$ is majority-consistent. If $\alpha \notin M(P'^{\downarrow Y})$ then *a fortiori* $\alpha \notin M(P^{\downarrow Y})$, which contradicts (4.22). Therefore, $\alpha \in M(P'^{\downarrow Y})$, and because of (4.22), it is also a maximal majority-consistent sub-agenda for P. We have shown that the maximal majority-consistent sub-agendas for P and P' coincide, therefore $R_{MSA}(P) = R_{MSA}(P')$. The proof for R_{MCSA} can be generated in exactly the same way. □

Proposition 4.8.2. *R_{RA} satisfies (IR-s).*

Proof. Let $\alpha \in \overline{\mathcal{A}}$ and assume that $\alpha \in T_{R_{RA}}(P)$. Then all sub-agendas in $R_{RA}(P)$ contain α. Let P' be an α-improvement of P. Then $N(P', \alpha) > N(P, \alpha)$, $N(P', \neg\alpha) < N(P, \neg\alpha)$, whereas for all $\varphi \neq \alpha, \neg\alpha$, $N(P', \varphi) = N(P, \varphi)$. Note that in $\geqslant_{P'}$, α appears either at an earlier position or in the same position as in $\geqslant_P$. Therefore, if $\succ'$ is an order refining $\geqslant_{P'}$, when α is considered in $\succ$, otherwise there would be an order $\succ$ refining $\geqslant_P$ resulting in a sub-agenda not containing α. Therefore α belongs to all sub-agendas in $RA(P')$.

□

Proposition 4.8.3. *R_{MWA} i.e., $\Delta^{d_H,\Sigma}$, satisfies (IR-s).*

Proof. Let P be a profile $P = (\hat{A}_1, \ldots, \hat{A}_k, \ldots, \hat{A}_n)$. Let the α-reinforcement of P be a profile $P' = (\hat{A}'_1, \ldots, \hat{A}'_k, \ldots, \hat{A}'_n) = (\hat{A}_1, \ldots, \hat{A}^*_k, \ldots, \hat{A}_n)$. Let $\hat{\mathbb{A}}$ be the set of all consistent and complete judgment sets over an agenda $\mathcal{A}$. Let us define $D(\hat{A}, P) = \sum_{i=1}^{n} d_H(A, A_i)$.

We have the following assumptions:

- $\alpha \notin \hat{A}_k$,
- $\alpha \in \hat{A}^*_k$,
- for all $\psi \in \overline{\mathcal{A}}, \psi \notin \{\alpha, \neg\alpha\}$ it holds $\psi \in \hat{A}_k$ iff $\psi \in \hat{A}^*_k$,
- $\alpha \in \hat{A}$, for all $\hat{A} \in R^{d_H,\Sigma}(P)$.

We first show that all the judgment sets $\hat{A} \in R^{d_h,\Sigma}(P)$ are such that $\hat{A} \in R^{d_h,\Sigma}(P')$ and that there exists no $\hat{A}' \in \hat{\mathbb{A}}$ such that: $\alpha \in \hat{A}'$, $\hat{A}' \notin R^{d_h,\Sigma}(P)$, but $\hat{A}' \in R^{d_h,\Sigma}(P')$.

Let the score of the winner judgment sets $\hat{A}$ for P be c, namely let $c = \sum_{i=1}^{n} d_H(A, A_i)$, for all $\hat{A} \in R^{d_h,\Sigma}(P)$. We have that, for all $\hat{A}' \in \hat{\mathbb{A}}$, when the cardinality of the pre-agenda is m:
$d_H(A', A_k) = m - |\hat{A}_k \cap \hat{A}'|$,
$d_H(A', A^*_k) = m - |\hat{A}^*_k \cap \hat{A}'|$.

Let $\hat{A}' \in \hat{\mathbb{A}}$ be such that $\alpha \in \hat{A}'$. Since $d_H(A_k, A^*_k) = 1$, we have that $|\hat{A}_k \cap \hat{A}'| - |\hat{A}^*_k \cap \hat{A}'| = 1$. Hence $d_H(A', A_k) = 1 + d_H(A', A^*_k)$ and

$$D(\hat{A}', P) = 1 + D(\hat{A}', P'). \tag{4.23}$$

For all the winners $\hat{A}$ for P, we obtain that $D(\hat{A}, P) = 1 + D(\hat{A}, P')$, hence

$$D(\hat{A}, P') = c - 1. \tag{4.24}$$

If an $\hat{A}' \notin R^{d_h,\Sigma}(P)$, then $D(\hat{A}', P) > c$ and due to 4.23, $D(\hat{A}', P') > c - 1$. We can conclude that there is no $\hat{A}' \in \hat{\mathbb{A}}$ such that $\alpha \in \hat{A}'$ and $\hat{A}' \notin R^{d_h,\Sigma}(P)$ but $\hat{A}' \in R^{d_h,\Sigma}(P')$.

We now show that there exists no $\hat{A}'' \in \hat{\mathbb{A}}$ such that $\alpha \notin \hat{A}''$ and $\hat{A}'' \notin R^{d_h,\Sigma}(P)$ but $\hat{A}'' \in R^{d_h,\Sigma}(P')$. We construct a proof by contradiction, starting with the assumption that there exists such a $\hat{A}'' \in R^{d_h,\Sigma}(P')$.

Since $\hat{A}'' \notin R^{d_h,\Sigma}(P)$, we obtain

$$D(\hat{A}'', P) > c. \tag{4.25}$$

Since $\alpha \in \hat{A}_k^*$ and $\alpha \notin \hat{A}'$, we obtain $d_H(A'', A_k) < d_H(A'', A_k^*)$ and consequently

$$D(\hat{A}'', P) < D(\hat{A}'', P'). \tag{4.26}$$

Putting together inequalities 4.25 and 4.26 we obtain

$$D(\hat{A}'', P') > c \tag{4.27}$$

However, the inequality 4.27 and inequality 4.24 are contradictory with the assumption that $\hat{A}'' \notin R^{d_h,\Sigma}(P)$. This completes the proof that $R^{d_h,\Sigma}$ is insensitive to reinforcement of collective judgements.

□

Proposition 4.8.4. *R_Y does not satisfy (IR-s).*

Voters	p	q	$p \wedge q$	r
2×	+	+	+	+
2×	+	-	-	+
1×	+	-	-	+
4×	-	+	-	-
$M(P)$	+	+	-	+

Voters	p	q	$p \wedge q$	r
2×	+	+	+	+
2×	+	-	-	+
1×	-	-	-	+
4×	-	+	-	-
$M(P)$	-	+	-	+

Table 4.5: P on the left, and the $\neg p$-reinforcement P' on the right used to show that R_Y does not satisfy (IR-s).

Proof. We use a proof by counter-example. Let the agenda be $\mathcal{A} = \{p, q, p \wedge q, r\}$. Consider the profile P in Table 4.5. P is not majority-consistent, but removing any voter who has p in her judgment set suffices to restore consistency, therefore $R_Y(P) = \{q, \neg(p \wedge q)\}$.

Consider the $\neg p$-reinforcement profile P', Table 4.5 right-most. $R_Y(P') = \{\neg p, q, \neg(p \wedge q), r\}$. Observe that although $\neg p \in T_Y(P)$ and $\neg p \in T_Y(P')$, $R_Y(P) \neq_T R_Y(P')$.

□

Proposition 4.8.5. *R_{RY} does not satisfy (IR-s).*

Proof. Consider again the profile P given on Table 4.5. We obtain $R_{RY}(P)$ by adding the fourth judgment set once, so that it appears five times in the profile instead of four. $R_{RY}(P) = \{q, \neg(p \wedge q)\}$, hence $T_{R_{RY}}(P) \models \neg p$ and $T_{R_{RY}}(P) \not\models t$. Consider the $\neg p$-reinforcement P' given on Table 4.5. $R_{RY}(P') = M(P') = \{\neg p, q, \neg(p \wedge q), t\}$. It can be observed that $R_{RY}(P) \neq R_{RY}(P')$. □

Proposition 4.8.6. *R_{MNAC} does not satisfy (IR-s).*

$Voters$	p	q	$p\wedge q$	$p\wedge r$	$p\wedge s$
$1\times\hat{A}_1$	+	+	+	−	−
$1\times A_2$	+	−	−	+	−
$1\times\hat{A}_3$	−	+	−	−	+
$M(P)$	+	+	−	−	−

$Voters$	p	q	$p\wedge q$	$p\wedge r$	$p\wedge s$
$1\times\hat{A}_1$	+	+	+	−	−
$1\times A_2$	+	−	−	−	−
$1\times\hat{A}_3$	−	+	−	−	+
$M(P')$	+	+	−	−	−

Table 4.6: The profiles P (left) and P' (right), an example that R_{MNAC} is not insensitive to reinforcement of collective judgements.

Proof. Consider the agenda $\mathcal{A}=\{p,q,p\wedge q,p\wedge r,q\wedge s\}$ and the profile P given in Table 4.6.

There are 6 profiles P_i such that $D(P,P_i)=2$, see Table 4.7. $R_{MNAC}(P)=\{\{p,q,p\wedge q,\neg(p\wedge r),\neg(p\wedge s)\},\{p,\neg q,\neg(p\wedge q),\neg(p\wedge r),\neg(p\wedge s)\},\{\neg p,q,\neg(p\wedge q),\neg(p\wedge r),\neg(p\wedge s)\}\}$. We have that $T_{R_{MNAC}}(P)\models\neg(p\wedge r)$.

P' is a $\neg(p\wedge r)$-reinforcement of P, but $R_{MNAC}(P')=\{\{\neg p,q,\neg(p\wedge q),\neg(p\wedge r),\neg(p\wedge s)\}\}$, since $D(P',P_3)=1$.

$Voters$	p	q	$p\wedge q$	$p\wedge r$	$p\wedge s$
$1\times$	+	−	−	−	−
$1\times$	+	−	−	+	−
$1\times$	−	+	−	−	+
$M(P_1)$	+	−	−	−	−

$Voters$	p	q	$p\wedge q$	$p\wedge r$	$p\wedge s$
$1\times$	+	+	+	−	−
$1\times$	+	+	+	+	−
$1\times$	−	+	−	−	+
$M(P_4)$	+	+	+	−	−

$Voters$	p	q	$p\wedge q$	$p\wedge r$	$p\wedge s$
$1\times$	−	+	−	−	−
$1\times$	+	−	−	+	−
$1\times$	−	+	−	−	+
$M(P_2)$	−	+	−	−	−

$Voters$	p	q	$p\wedge q$	$p\wedge r$	$p\wedge s$
$1\times$	+	+	+	−	−
$1\times$	+	−	−	+	−
$1\times$	−	−	−	−	−
$M(P_5)$	+	−	−	−	−

$Voters$	p	q	$p\wedge q$	$p\wedge r$	$p\wedge s$
$1\times$	+	+	+	−	−
$1\times$	−	−	−	−	−
$1\times$	−	+	−	−	+
$M(P_3)$	−	+	−	−	−

$Voters$	p	q	$p\wedge q$	$p\wedge r$	$p\wedge s$
$1\times$	+	+	+	−	−
$1\times$	+	−	−	+	−
$1\times$	+	+	+	−	+
$M(P_6)$	+	+	+	−	−

Table 4.7: The profiles P_i, $i\in[1,6]$ for which $D(P,P_i)=2$. Note that $D(P',P_3)=1$.

□

Proposition 4.8.7. *$R^{d_H,max}$ does not satisfy (IR-s).*

Proof. Consider the agenda $\mathcal{A}=\{p\wedge r,p\wedge q,q,t\}$, the profile P for three agents:

Voters	$p \wedge r$	$p \wedge q$	q	t
$1\times$	$-$	$+$	$-$	$+$
$1\times$	$+$	$-$	$-$	$-$
$1\times$	$-$	$-$	$+$	$+$

and its $\neg(p \wedge q)$-reinforcement (in the first voter's judgment set) P':

voters	$p \wedge r$	$p \wedge q$	q	t
1	$-$	$-$	$-$	$+$
2	$+$	$-$	$-$	$-$
3	$-$	$-$	$+$	$+$

As it can be observed from Table 4.8, $\neg(p \wedge q) \in T_{R^{d_H,max}}(P)$, since $R^{d_H,max}(P) = \{\{\neg(p \wedge r), \neg(p \wedge q), \neg q, \neg t\}, \{\neg(p \wedge r), \neg(p \wedge q), \neg q, t\}, \{p \wedge r, \neg(p \wedge q), \neg q, t\}\}$. However, as it can be observed from Table 4.9, $R^{d_H,max}(P') = \{\{\neg(p \wedge r), \neg(p \wedge q), \neg q, \neg t\}, \{\neg(p \wedge r), \neg(p \wedge q), \neg q, t\}, \{p \wedge r, \neg(p \wedge q), \neg q, t\}, \{\neg(p \wedge r), \neg(p \wedge q), q, \neg t\}\}$. Thus, $\neg(p \wedge q) \in T_{R^{d_H,max}}(P')$ and $R^{d_H,max}(P) \neq R^{d_H,max}(P')$. Furthermore, since $\neg q \in T_{R^{d_H,max}}(P)$, but $\neg q \notin T_{R^{d_H,max}}(P')$, we obtain that $R^{d_H,max}(P) \neq_T R^{d_H,max}(P')$.

$\hat{A} \in \hat{\mathbb{A}}$		$d_H(\hat{A},\hat{A}_5)$	$d_H(\hat{A},\hat{A}_8)$	$d_H(\hat{A},\hat{A}_4)$	max
$\hat{A}_1$	$\{\neg(p \wedge r), \neg(p \wedge q), \neg q, \neg t\}$	2	1	2	2
$\hat{A}_2$	$\{\neg(p \wedge r), \neg(p \wedge q), \neg q, t\}$	1	2	1	2
$\hat{A}_3$	$\{\neg(p \wedge r), \neg(p \wedge q), q, \neg t\}$	3	2	1	3
$\hat{A}_4$	$\{\neg(p \wedge r), \neg(p \wedge q), q, t\}$	2	3	0	3
$\hat{A}_5$	$\{\neg(p \wedge r), p \wedge q, \neg q, t\}$	0	3	2	3
$\hat{A}_6$	$\{\neg(p \wedge r), p \wedge q, q, \neg t\}$	2	3	2	3
$\hat{A}_7$	$\{\neg(p \wedge r), p \wedge q, q, t\}$	1	4	1	4
$\hat{A}_8$	$\{p \wedge r, \neg(p \wedge q), \neg q, \neg t\}$	3	0	3	3
$\hat{A}_9$	$\{p \wedge r, \neg(p \wedge q), \neg q, t\}$	2	1	2	2
$\hat{A}_{10}$	$\{p \wedge r, p \wedge q, q, \neg t\}$	3	2	3	3
$\hat{A}_{11}$	$\{p \wedge r, p \wedge q, q, t\}$	2	3	2	3

Table 4.8: The max of Hamming metrics from an element in the set $\hat{\mathbb{A}}$ to each of the agent's judgment sets in profile P.

□

We now consider the Δ rules.

Proposition 4.8.8. *If $d \in \{d_T, d_H\}$ then $\Delta^{d,\odot}$ satisfies (IR).*

Proof. Consider a profile π for $\mathcal{A}$, a $\beta \in \overline{\mathcal{A}}$, a profile π' which is a β-reinforcement of π. The k-th row is the row on which π and π' differ. Observe that, for $\beta \in \{b, \neg b\}$ or $b = \neg\beta$ $(\pi' \triangleright k) \triangledown b = v(b)$, $v(b) \in \{0,1\}$ and $(\pi \triangleright k) \triangledown b \in \{0, \frac{1}{2}, 1\}$. Let $A \in \Delta^{d,\odot}(\pi, W)$, $A \triangledown b = v(b)$. We need to show that all $A' \in \Delta^{d,\odot}(\pi', W)$ are such that $A' \triangledown b = v(b)$. We construct a proof by contradiction.

$\hat{A} \in \hat{\mathbb{A}}$		$d_H(\hat{A},\hat{A}_2)$	$d_H(\hat{A},\hat{A}_8)$	$d_H(\hat{A},\hat{A}_4)$	max
$\hat{A}_1$	$\{\neg(p \wedge r), \neg(p \wedge q), \neg q, \neg t\}$	2	1	2	2
$\hat{A}_2$	$\{\neg(p \wedge r), \neg(p \wedge q), \neg q, t\}$	1	2	1	2
$\hat{A}_3$	$\{\neg(p \wedge r), \neg(p \wedge q), q, \neg t\}$	1	2	1	2
$\hat{A}_4$	$\{\neg(p \wedge r), \neg(p \wedge q), q, t\}$	0	3	0	3
$\hat{A}_5$	$\{\neg(p \wedge r), p \wedge q, \neg q, t\}$	2	3	2	3
$\hat{A}_6$	$\{\neg(p \wedge r), p \wedge q, q, \neg t\}$	1	3	2	3
$\hat{A}_7$	$\{\neg(p \wedge r), p \wedge q, q, t\}$	1	4	1	4
$\hat{A}_8$	$\{p \wedge r, \neg(p \wedge q), \neg q, \neg t\}$	3	0	3	3
$\hat{A}_9$	$\{p \wedge r, \neg(p \wedge q), \neg q, t\}$	2	1	2	2
$\hat{A}_{10}$	$\{p \wedge r, p \wedge q, q, \neg t\}$	2	2	3	3
$\hat{A}_{11}$	$\{p \wedge r, p \wedge q, q, t\}$	1	3	2	3

Table 4.9: The max of Hamming metrics from an element in the set $\hat{\mathbb{A}}$ to each of the agent's judgment sets in profile P'.

We use the notation

$$d_w(A,A_i) = \sum_{j=1}^{m} w(i,j) \cdot \delta(A \triangledown a_j, A_i \triangledown a_j),$$

where $\delta \in \{\delta_H, \delta_T\}$.

From $A \in \Delta^{d,\odot}(\pi, W)$ it follows that

$$\odot_{i=1}^{n} d_w(A, \pi \triangleright i) < \odot_{i=1}^{n} d_w(A', \pi \triangleright i). \tag{4.28}$$

Since $d_w(A, \pi \triangleright k) \geqslant d_w(A, \pi' \triangleright k)$ it follows, from the non-decreasingness of $\odot$:

$$\odot_{i=1}^{n} d_w(A, \pi \triangleright i) \geqslant \odot_{i=1}^{n} d_w(A, \pi' \triangleright i). \tag{4.29}$$

It follows from (4.28) and (4.29) that:

$$\odot_{i=1}^{n} d_w(A, \pi' \triangleright i) < \odot_{i=1}^{n} d_w(A', \pi \triangleright i). \tag{4.30}$$

Assume that there is a $A' \in \Delta^{d,\odot}(\pi', W)$ such that $A' \triangledown b \neq v(b)$, and as such $A' \notin \Delta^{d,\odot}(\pi, W)$. It follows that

$$\odot_{i=1}^{n} d_w(A', \pi' \triangleright i) < \odot_{i=1}^{n} d_w(A, \pi' \triangleright i). \tag{4.31}$$

From (4.30) and (4.31), it follows that

$$\odot_{i=1}^{n} d_w(A', \pi' \triangleright i) < \odot_{i=1}^{n} d_w(A', \pi \triangleright i). \tag{4.32}$$

From (4.32), since $\pi' \triangleright i = \pi \triangleright i$, for all $i \neq k$, it follows that

$$d_w(A', \pi' \triangleright k) < d_w(A', \pi \triangleright k). \tag{4.33}$$

Recall that

$$d_w(A', \pi' \triangleright k) = \sum_{j=1}^{m} w(k, j) \cdot \delta(A' \triangledown a_j, \pi'_{k,j})$$

and

$$d_w(A', \pi \triangleright k) = \sum_{j=1}^{m} w(k, j) \cdot \delta(A' \triangledown a_j, \pi_{k,j})$$

Since $\pi'_{k,j} = p_{k,j}$, for all $a_j \neq b$ It follows that (4.33) holds, if and only if

$$\delta(A' \triangledown b, \pi'_{k,b}) < \delta(A' \triangledown b, \pi_{k,b}) \quad (4.34)$$

For δ_H, the inequality (4.34) holds only when $\delta(A' \triangledown b, \pi'_{k,b}) = 0$, however that would imply that $A' \triangledown b = \pi'_{k,b}$ which is a contradiction with the assumption that $A' \triangledown b \neq v(b)$.

For δ_T, the inequality (4.34) holds only when it is possible that $\delta(A' \triangledown b, \pi'_{k,b}) = \frac{1}{2}$ or $\delta(A' \triangledown b, \pi_{k,b}) = 1$. If $\delta(A' \triangledown b, \pi'_{k,b}) = \frac{1}{2}$, then $A' \triangledown b = \frac{1}{2}$ due to $v(b) \in \{0, 1\}$. However, if $A' \triangledown b = \frac{1}{2}$, then $\delta(A' \triangledown b, (\pi \triangleright k) \triangledown b) \in \{0, \frac{1}{2}\}$ and we reach a contradiction again. □

Proposition 4.8.9. $\Delta^{d_D, \odot}$ *satisfies (IR).*

Proof. We make the same assumptions as in Proposition 4.8.8. Consider a profile π for $\mathcal{A}$, a $\beta \in \overline{\mathcal{A}}$, a profile π' which is a β-reinforcement of π. The k-th row is the row on which π and π' differ. Observe that, for $\beta \in \{b, \neg b\}$ or $b = \neg\beta$ $(\pi' \triangleright k) \triangledown b = v(b)$, $v(b) \in \{0, 1\}$ and $(\pi \triangleright k) \triangledown b \in \{0, \frac{1}{2}, 1\}$. Let $A \in \Delta^{d_D, \odot}(\pi, W)$, $A \triangledown b = v(b)$. We need to show that all $A' \in \Delta^{d_D, \odot}(\pi', W)$ are such that $A' \triangledown b = v(b)$. We construct a proof by contradiction.

We use the notation

$$d_w(A, A_i) = max(w_{i,1} \cdot \delta_h(A(1), A_i(1)), \ldots, w_{i,m} \cdot \delta_h(A(m), A_i(m))).$$

From $A \in \Delta^{d_D, \odot}(\pi)$ it follows that

$$\odot_{i=1}^{n} d_w(A, \pi \triangleright i) < \odot_{i=1}^{n} d_w(A', \pi \triangledown i). \quad (4.35)$$

Since $d_D(A, \pi \triangledown k) \geqslant d_D(A, \pi' \triangledown k)$ it follows, from the non-decreasingness of $\odot$:

$$\odot_{i=1}^{n} d_w(A, \pi \triangledown i) \geqslant \odot_{i=1}^{n} d_w(A, \pi' \triangledown i). \quad (4.36)$$

It follows from (4.35) and (4.36) that:

$$\odot_{i=1}^{n} d_w(A, \pi' \triangledown i) < \odot_{i=1}^{n} d_w(A', \pi \triangledown i). \quad (4.37)$$

Assume that $A' \in \Delta^{d_D, \odot}(\pi', W)$ and $A' \notin \Delta^{d_D, \odot}(\pi, W)$. It follows that

$$\odot_{i=1}^{n} d_w(A', \pi' \triangledown i) < \odot_{i=1}^{n} d_w(A, \pi' \triangledown i). \quad (4.38)$$

From (4.37) and (4.38), it follows that

$$\odot_{i=1}^{n} d_w(A', \pi' \triangledown i) < \odot_{i=1}^{n} d_w(A', \pi \triangledown i). \quad (4.39)$$

From (4.39), since $\pi' \triangledown i = \pi \triangledown i$, for all $i \neq k$, it follows that

$$d_D(A', \pi' \triangledown k) < d_D(A', \pi \triangledown k). \quad (4.40)$$

The inequality (4.40) is only possible when $d_D(A', \pi' \triangledown k) = 0$, however that would imply that $A' \triangledown b = (\pi' \triangleright k) \triangledown b$ which is a contradiction with the assumption that $A' \triangledown b \neq v(b)$. □

4.9 Separability

In addition to the relational properties that are considered in judgment aggregation, and which we lifted in the previous sections, we can also introduce relational properties inspired by properties of interest studied in voting theory.

In voting theory, the separability property states that if an alternative is a winner under a voting rule, for two distinct profiles under the same set of candidates, then that alternative is a winner, under the same voting rule, for the profile obtained by combining the two profiles. The property of separability is defined in (Smith, 1973), also defined as consistency in (Young, 1975), and it is sometimes called reinforcement as well. This property is best known as one of the conditions, together with neutrality and anonymity, used by Young in his characterization of scoring social choice rules (Young, 1975). The voting rules that do not satisfy the separability property are subject to occurrences of *Simpson's paradox* (Blyth, 1972).

In judgment aggregation, the separability property is of interest as well. One reason is that the separability property is a natural requirement to make: if a judgment set is among the collective judgment sets for profile π_1 and for profile π_2, then it should be among the judgment sets for the combined profile P.

Since the judgment aggregation sequence can be considered as a solid piece of information or as divisible collection of judgments, we can define at least two versions of a separability property in judgment aggregation: sequence-separability (S-s) and issue-separability (S-i).

Definition 72 (Horizontal merge). *Let M_1 be a $n_1 \times m$ matrix and M_2 a $n_2 \times m$ matrix. The matrix M is called a horizontal merge of M_1 and M_2 if $M \triangleright i = M_1 \triangleright i$ for all $i \in [1, n_1]$ and $M \triangleright (j + n_1) = M_1 \triangleright j$ for all $j \in [1, n_2]$.*

Definition 73 (S-i). *A rule F satisfies* issue-separability *when for every $\pi \in S_I^{n_1}$, $W_1 \in (\mathbb{R}^+)^{n_1 \times m}$ and $\pi \in S_I^{n_2}$, $W_2 \in (\mathbb{R}^+)^{n_2 \times m}$ and their horizontal merge $\pi \in S_I^{n_1+n_2}$, $W \in (\mathbb{R}^+)^{(n_1+n_2) \times m}$ if [for all $A \in F(\pi_1, W_1)$, $A(j) = \alpha$] and [for all $A' \in F(\pi_2, W_2)$, $A'(j) = \alpha$], then [for all $A'' \in F(\pi, W)$, $A''(j) = \alpha$].*

Definition 74 (S-s). *A rule F satisfies* sequence-separability *when for every $\pi \in S_I^{n_1}$, $W_1 \in (\mathbb{R}^+)^{n_1 \times m}$ and $\pi \in S_I^{n_2}$, $W_2 \in (\mathbb{R}^+)^{n_2 \times m}$ and their horizontal merge $\pi \in S_I^{n_1+n_2}$, $W \in (\mathbb{R}^+)^{(n_1+n_2) \times m}$ if $F(\pi_1, W_1) \cap F(\pi_2, W_2) \neq \emptyset$, then $F(\pi_1, W_1) \cap F(\pi_2, W_2) \subseteq F(\pi, W)$.*

The issue-separability states that if a judgment is in the theory of the rule on a profile π_1 and it is in the theory of profile π_2, both being profiles on the same agenda and constraints, then the same judgment is in the theory of the horizontal merge of π_1 and π_2. The sequence-separability looks at whole sequences instead of judgments. If a sequence is selected as collective by a rule for π_1 and the same sequence is selected as collective for π_2, then that sequence is among the collective sequence selected by the rule for the horizontal merge of π_1 and π_2.

Unlike with unanimity and unanimity principle, majoritarianism and majority-preservation, and (IR-s) and (IR), one of the separability properties does not imply the other. We study issue-separability for the rules based on minimization (again not considering R_{IY} and R_{MR}) and sequence-separability for the weighted distance-based rules.

For the rules in Chapter 2 we study (S-i) by establishing a general result[3] which shows that majority-preservation and rule (S-i) are incompatible. This result can be seen as the judgment aggregation counterpart of the result that states that every Condorcet-consistent voting rule violates reinforcement, see Theorem 9.2 (Moulin, 1991, pg.237).

Proposition 4.9.1. *If a rule aggregator is majority-preserving then it violates issue-separability.*

Proof. Let R be a majority-preserving rule, and assume furthermore that R satisfies (S-i). Let $\mathcal{A} = \{p, q, p \vee q\}$, and P the 10-voter profile as follows:

voters	p	q	$p \vee q$
1,2	+	+	+
3,4	−	+	+
5,6	+	−	+
7,8,9,10	−	−	−

Consider the two sub-profiles P_1 consisting of voters $\{1,3,4,7,8\}$ and P_2 consisting of voters $\{2,5,6,9,10\}$. P_1 and P_2 are majority-consistent, with $M(P_1) = \{\neg p, q, p \vee q\}$ and $M(P_2) = \{p, \neg q, p \vee q\}$. Since R is majority-preserving, we have $R(P_1) = \{\{p, \neg q, p \vee q\}\}$ and $R(P_2) = \{\{q, \neg p, p \vee q\}\}$; therefore, $p \leftrightarrow \neg q \in T_R(P_1)$ and $p \leftrightarrow \neg q \in T_R(P_2)$, from which, by (S-i),

$$p \leftrightarrow \neg q \in T_R(P_1 \cup P_2) = T_R(P). \tag{4.41}$$

Consider now the two sub-profiles P_3 consisting of voters $\{1,2,3\}$ and P_4 consisting of voters 4 to 10. The profiles P_3 and P_4 are majority-consistent, with $M(P_3) = \{p, q, p \wedge q\}$ and $M(P_4) = \{\neg p, \neg q, \neg p \wedge \neg q\}$. Since R is majority-preserving, we have $R(P_3) = \{p, q, p \wedge q\}$ and $R(P_4) = \{\neg p, \neg q, \neg p \wedge \neg q\}$. As a consequence

$$p \leftrightarrow q \in T_R(P_3) \text{ and } p \leftrightarrow \neg q \in T_R(P_4). \tag{4.42}$$

From (4.42), by separability,

$$p \leftrightarrow q \in T_R(P_1 \cup P_2) = T_R(P). \tag{4.43}$$

The equation (4.43) is in contradiction with (4.41). □

As a corollary, all the rules based on minimization except $R^{d_H,\max}$ violate issue-separability.

Corollary 4.9.2. *The aggregation rules R_Y, R_{MSA}, R_{MCSA}, R_{MWA}, R_{RA}, R_{RY}, and R_{MNAC} do not satisfy (S-i).*

[3] This theorem was proved by Jérôme Lang.

The only one of our rules which is not majority-preserving is $R^{d_H,\max}$. However, this one does not satisfy (S-i) either, which shows that it seems extremely difficult to find a reasonable judgment aggregation rule that satisfies (S-i).

Proposition 4.9.3. *$R^{d_H,\max}$ does not satisfy (S-i).*

Proof. Let $\mathcal{A} = \{p, q, r, p \to (q \wedge r)\}$, and the 5-voter profile P:

voters	p	q	r	$p \to (q \wedge r)$
1,2,3	+	+	−	−
4,5	+	+	+	+

Consider also the two sub-profiles P_1 consisting of voters 1, 2 and 3, and P_2 consisting of voters 4 and 5. Observe that $R^{d_H,max}(P_1) = \{\{p, q, \neg r, \neg(p \to (q \wedge r))\}\}$ and $R^{d_H,max}(P_2) = \{\{p, q, r, p \to (q \wedge r)\}\}$, thus $p \in T_{R^{d_H,max}}(P_1)$ and $p \in T_{R^{d_H,max}}(P_2)$. However, $R^{d_H,max}(P) = \{\{p, q, \neg r, \neg(p \to (q \wedge r))\}, \{p, q, r, p \to (q \wedge r)\}, \{\neg p, q, \neg r, p \to (q \wedge r)\}, \{p, \neg q, r, \neg(p \to (q \wedge r))\}\}$, therefore $p \notin T_{R^{d_H,max}}(P)$. □

Regarding sequence-separability and the weighted distance-based rules, there are more positive results.

Proposition 4.9.4. *If $\odot$ is associative (see Definition 31), then $\Delta^{d,\odot}$ satisfies (S-s).*

Proof. Let $s(A, A_i) = \circledast_{j=1}^{m} w(i, j)\delta(A(j), A_i(j))$. Due to associativity

$$\odot(\underbrace{s(A, A_1), \ldots, s(A, A_{k_1})}_{\mathbf{x}(A)}, \underbrace{s(A, A_{k_2}), \ldots, s(A, A_n)}_{\mathbf{y}(A)}) = \odot(\odot(\mathbf{x}(A)), \odot(\mathbf{y}(A))).$$

Due to the non-decreasing of $\odot$, if

$$\mathbf{x}(A) \leqslant \mathbf{x}(A') \text{ and } \mathbf{y}(A) \leqslant \mathbf{y}(A'), \tag{4.44}$$

then $\odot(\mathbf{x}(A), \mathbf{y}(A)) \leqslant \odot(\mathbf{x}(A'), \mathbf{y}, (A'))$.

Let $\pi_1 \in \mathbb{A}^{n_1}$, $\pi_2 \in \mathbb{A}^{n_2}$, $W_1 \in (\mathbb{R}^+)^{n_1}$ and $W_2 \in (\mathbb{R}^+)^{n_2}$. If $A \in \Delta^{d,\odot}(\pi_1, W_1)$, $A \in \Delta^{d,\odot}(\pi_2, W_2)$, then (4.44) holds for each $A' \in \mathbb{A}$ and consequently $A \in \Delta^{d,\odot}(\pi, W)$ for the horizontal merges π and W. □

The function Π^* is not associative, however it satisfies sequence-separability.

Proposition 4.9.5. *Δ^{d,Π^*} satisfies (S-s).*

Proof. The proof can be constructed similarly as the proof of Theorem 4.9.4. Observe that

$$\prod(s(A, \pi \triangleright 1) + 1, \ldots, s(A, \pi \triangleright n_1) + 1, s(A, \pi \triangleright (n_1 + 1)) + 1, \ldots s(A, \pi \triangleright n_2) + 1)$$

can be written as

$$\prod(s(A, \pi_1 \triangleright 1) + 1, \ldots, s(A, \pi_1 \triangleright n_1) + 1) \cdot \prod(s(A, \pi_2 \triangleright 1) + 1, \ldots, s(A, \pi_2 \triangleright n_2) + 1). \tag{4.45}$$

For $x_1, x_2, y_1, y_2 \in \mathbb{R}^+$ if $x_1 < x_2$ and $y_1 < y_2$, then $x_1 \cdot y_1 \leqslant x_2 \cdot y_2$. Consequently, if $A \in \Delta^{d,\Pi^*}(\pi_1, W_1)$ and $A \in \Delta^{d,\Pi^*}(\pi_2, W_2)$, then $A \in \Delta^{d,\Pi^*}(\pi, W)$. □

4.10 Other properties for judgment aggregation rules

In addition to separability, we can construct other properties that might be desirable for an aggregation rule to satisfy, inspired by properties studied in voting theory. In this section we give their definitions.

Tideman (1987) introduced the *clone independence criterion* for voting rules. This property states that when a candidate is added to the set of candidates, and this candidate is identical to a candidate already in the set, the winner of the election will not change.

There are no candidates in judgment aggregation, but the independence of clones can be defined for agenda items. Indeed, if two agenda issues are logically equivalent, then the collective judgment on the both should be the same. An idea that the rule should be insensitive to agenda clones already appears in (Dietrich, 2006b), where a similar property is defined as *logical agenda manipulation*. The difference between our definition and the one in (Dietrich, 2006b) is that in (Dietrich, 2006b) one speaks of *settled issues*, namely issues whose truth-value is determined by any judgment set (consistent and complete).

Cariani et al. (2008) define a property for function aggregators called *translation invariance*. A function aggregator is translation invariant if the collective judgment set does not depend on the particular language used to model the agenda. Namely, if two agendas are semantically equivalent and two equal profiles, each for one of the agendas, are aggregated, then the collective judgment sets selected for each profile should be the same. Cariani et al. (2008) prove that whether a function is translation invariant depends on the atoms in the agenda.

We define clones as issues in the agenda that are logically equivalent. All clones are settled issues, but not all settled issues are clones.

Definition 75 (Clones). *Given an agenda $\mathcal{A}$, issues $a, a' \in \mathcal{A}$ are clones when for all $A \in S_I$, $A(a) = x$ if and only if $A(a') = x$.*

Let M be a $n \times m$ matrix and $Y \subset [1,m]$ a set of columns. The sub-matrix $M^{\downarrow Y}$ is the $n \times (m - |Y|)$ matrix obtained by removing the columns in Y from M. Consider as an example the 3×3 matrix M , $Y = \{2\}$ and $M^{\downarrow Y}$:

$$M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \qquad M^{\downarrow Y} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}$$

Definition 76 (IAC). *Let $\mathcal{A}$ be an agenda containing the clones a and a_k, and let $Y = \{k\}$. A judgment aggregation rule F is* independent of agenda clones *when for every profile $\pi \in S_I^n$, if $A \in F(\pi, U)$ and $A' \in F(p^{\downarrow Y}, W^{\downarrow Y})$, then for all $a_j \neq a_k$ $A(j) = A'(j)$.*

Example 4.10.1. *Let $\mathcal{A} = \{p, q, p \wedge q, p \wedge p\}$. We can observe that p and $p \wedge p$ are clones. Consider the profiles P and the reduced P' in Table 4.10. On this profile the rule R_Y is insensitive to clones, while the rule R_{MCSA} is not.*

There are aggregation contexts in which a group of agents needs to make judgments on the same agenda in various moments in time.

Example 4.10.2. *Consider a group of agents that develops and maintains software. The group needs to make decisions regarding the resources spent on developing and maintaining the software. Let the issues be:*

	p	q	$p \wedge q$	$p \wedge p$
	+	-	-	+
	-	+	-	-
	+	+	+	+
R_Y		+		
	+			+
			-	
R_{MCSA}	+	+		+
	+		-	+

	p	q	$p \wedge q$
	+	-	-
	-	+	-
	+	+	+
R_Y		+	
	+		
			-
R_{MCSA}	+	+	
	+		-
		+	-

Table 4.10: The profile P containing the judgments on the clones (left) and the sub-profile P' without the judgments on $p \wedge p$.

- *release new version (p),*
- *fix known bugs (q),*
- *improve the user interface (s) ,*
- *advertise the product (t).*

Let the relations between these issues be $p \rightarrow t$, $q \rightarrow p$ and $q \veebar s$[4].

As long as the software is maintained, the group would need to make decisions on the same or some of the agenda issues. At a given moment, after eliciting the group opinions on the full agenda, the group might need to use the decisions regarding p, q and t but determines that s is not really of interest.

In Example 4.10.2, how should a collective judgment on p change if the agenda, and profile, is reduced by s? It is intuitively undesirable that the decision on p should change since. We call this property *insensitivity to agenda shrinking*.

Definition 77 (IAS). *Let $\mathcal{A}$ be an agenda, $a_k \in \mathcal{A}$ and $Y = \{k\}$. A judgment aggregation rule F satisfies the* strong insensitivity to agenda shrinking *when for all $\pi \in S^n_I$, $W \in (\mathbb{R}^+)^{n \times m}$ if [for all $A \in F(\pi, W)$ and for all $k \neq j$, $A(j) = x$], then [for all $A' \in F(\pi^{\downarrow Y}, W^{\downarrow Y})$ and for all $k \neq j$, $A'(j) = x$].*

Example 4.10.3. *Consider $\mathcal{A} = \{p, q, q \rightarrow p, p \rightarrow t, s, q \veebar s\}$, the profile P for it, and the profile P' for $\mathcal{A}$ shrunk for s, given on Table 4.11. As we can observe from the table, $T_{R_{MCSA}}(P) \models p$ but $T_{R_{MCSA}}(P') \not\models p$, hence the rule R_{MCSA} does not satisfy the property.*

When a rule is based on minimization, we can expect that it will be sensitive to the shrinking of the agenda for an arbitrary issue, since the removed issue might share sub-formulas with other issues, as it was the case in Example 4.10.3. We can consider the insensitivity to agenda shrinking when the issues removed are atomic and not part of other issues. This property we call *insensitivity to atomic agenda shrinking*.

[4]The connective $\veebar$ is the exclusive or, defined as $\varphi \veebar \psi \equiv (\neg\varphi \wedge \psi) \vee (\varphi \wedge \neg\psi)$.

	p	q	$q \to p$	$p \to t$	s	$q \veebar s$
	-	-	+	+	+	+
	-	+	-	+	-	-
	+	+	+	-	-	+
R_Y		+			-	
			+			+
	-			+		
R_{MCSA}		+	+	+	-	+

	p	q	$q \to p$	$p \to t$	$q \veebar s$
	-	-	+	+	+
	-	+	-	+	-
	+	+	+	-	+
R_Y		+			
			+		+
	-			+	
R_{MCSA}		+	+	+	+
	-	+		+	+
	-		+	+	+

Table 4.11: The profile P containing the judgments on the full agenda and the sub-profile P' without the judgments on t.

Definition 78 (IAAS). *Let $\mathcal{A}$ be an agenda and let $Y \subset \mathcal{A}$ be the set of the indexes of each atomic formula $p \in \mathcal{A}$ that is not a sub-formula for any formula $\varphi \in \mathcal{A}\backslash Y$. A judgment aggregation rule F is insensitive to atomic agenda shrinking when for all $\pi \in §_I^n$, $W \in (\mathbb{R}^+)^{n\times m}$ if [for all $A \in F(\pi, W)$ and for all $k \neq j$, $A(j) = x$], then [for all $A' \in F(\pi^{\downarrow Y}, W^{\downarrow Y})$ and for all $k \neq j$, $A'(j) = x$].*

Another aggregation property regarding the agenda is the *agenda separation*. We first define what it means for an issue to be independent in an agenda.

Definition 79 (Independence of an issue)**.** *An issue a is* independent *in an agenda $\mathcal{A}$, when for any consistent judgment set $A \subset \overline{\mathcal{A}}$ and $\hat{a} \in \{a, \neg a\}$, $A\backslash\{a\} \not\models a$ and $A\backslash\{a\} \not\models \neg a$.*

Another way to define independent agenda issues is as conflict-free: a does not belong to any minimal inconsistent subset of $\mathcal{A}$. When an agenda contains an independent issue a, then it is reasonable to expect that, regardless of what the judgments on the other agenda issues are, the collective judgment on a always coincides with the majority of the judgments regarding a.

Definition 80 (AS)**.** *Let $\mathcal{A}$ be an agenda, and $a_j \in \mathcal{A}$ an independent issue. A judgment aggregation rule F satisfies* agenda separation *when for every $\pi \in S_I^n$, $W \in (\mathbb{R}^+)^{n\times m}$ if for all $A \in F(\pi, W)$, $A(j) = m(\pi \triangledown j, W)$.*

Example 4.10.4. *Consider the agenda $\mathcal{A} = \{p, q, p \wedge q, r\}$ and the profile for this agenda given in Table 4.12.*

Although this example satisfies the agenda separability property for both the R_Y and R_{MNAC} rules, this is not always the case. Finding a counter-example for R_Y is fairly simple.

The agenda separation property is desirable since it allows for profiles to be split into a part of judgments on dependent and part of judgments on independent issues, with only the judgments on dependent issues to be aggregated.

One property considered in voting theory is the property of *invulnerability to a no-show paradox*. Fishburn and Brams (1983) define the no-show paradox as “the addition of identical ballots with candidate x ranked last may change the winner from another candidate to x”. The

	p	q	$p \wedge q$	r
	+	-	-	+
	-	+	-	+
	+	+	+	-
R_Y		+		
	+			
			-	+
R_{MNAC}	-	+	-	+
	+	-	-	+

Table 4.12: The profile P containing the judgments on $\mathcal{A} = \{p, q, p \wedge q, r\}$ and the corresponding outputs from R_Y and R_{MNAC}.

no-show paradox occurs when a group of voters is better off by not voting than by voting according to its preferences (Nurmi, 2004). Moulin (1988) show that when there are at least four options (candidates), every voting rule that elects the Condorcet winner must generate the no-show paradox.

Are judgment aggregation rules susceptible to the no-show paradox as well? We can define when a rule is invulnerable by the no-show paradox.

Definition 81 (INS). *A judgment aggregation rule F satisfies the invulnerability to the no-show paradox when, for every $\pi \in S_I^n$, $W \in (\mathbb{R}^+)^{n \times m}$ and $A^* \in S_I$, $V \in (\mathbb{R}^+)^m$ such that $A^*(j) = x$ if for all $A \in F(\pi, W)$, $A(j) = x$, then for all $A' \in F(\pi', W')$. The profile π' and weights W' are the horizontal merge of π with A^* and W with V correspondingly.*

Example 4.10.5. *Consider the agenda $\mathcal{A} = \{p, q, p \wedge q\}$ and the profiles P and P' given in Table 4.13. We can observe that $\neg(p \wedge q) \in T_{R_{MNAC}}(P)$ and $\neg(p \wedge q) \in T_{R_{MNAC}}(P')$, hence for this profile R_{MNAC} is invulnerable to the no-show paradox.*

	p	q	$p \wedge q$
	+	-	-
	-	+	-
	+	+	+
R_{MNAC}	+	-	-
	-	+	-

	p	q	$p \wedge q$
	+	-	-
	-	+	-
	+	+	+
	-	-	-
R_{MNAC}			-

Table 4.13: The profile P (left) and an its extension P' (right) with $\{\neg p, \neg q, \neg(p \wedge q)\}$.

4.11 Conclusions

In this chapter designed rule aggregator properties and analyzed, with respect to these properties, the rules we defined in Chapters 2 and 3. We intend to use these properties to distinguish among the rules from Chapters 2 and 3 correspondingly.

Properties in judgment aggregation have been defined for function aggregators and binary unweight judgments. We first defined the correspondence between, on one side, a property defined for weighted rule aggregators and ternary judgments, and on the other side, a property defined for an unweight binary function aggregator. Due to the irresoluteness of the rule aggregators, there are multiple rule aggregator properties correspond to each function aggregator property.

There are not many properties considered in judgment aggregation theory. For each of the known property we defined at least one corresponding rule property and analyzed which of our rules satisfies it. The following rule aggregation properties have been included: universal domain, anonymity, neutrality, independence of irrelevant information, collective rationality, majority-preservation, majoritarianism, unanimity, unanimity preservation and monotonicity. In addition we considered separability, a rule aggregation property corresponding to a (separability) property in voting theory.

All the families of judgment aggregation rules we introduced satisfy the structural properties of universal domain, anonymity, neutrality, and collective rationality. None of the introduced rules satisfies the independence of irrelevant alternatives (in any of the defined versions). The relational properties we considered were majority adherence properties, unanimity properties, monotonicity properties and separability properties. The results are summarized in Table 4.14.

	Majority Preservation	Weak Unanimity	Strong Unanimity	IR-s	S-i
R_Y	✓	✓	✓	no	no
R_{MSA}	✓	✓	no	✓	no
R_{MCSA}	✓	no	no	✓	no
R_{MWA}	✓	no	no	✓	no
R_{RA}	✓	✓	✓	✓	no
$R^{d_H,max}$	no	no	no	no	no
R_{RY}	✓	✓	✓	no	no
R_{MNAC}	✓	no	no	no	no

Table 4.14: Summary of the results for the social theoretic properties of the judgment aggregation rules.

We considered the same properties for the examples of distances and aggregation functions we introduced. In Table 4.15 we summarize the results. With d, when no index is specified, we denote any product metric and with $\odot$ any aggregation function.

Whether a property holds for a distance-based rule sometimes depends on the properties of the chosen d and $\odot$, as was the case with anonymity and separability. On the other hand, *e.g.*, whether unanimity-principle is satisfied depends on the ratio between the cardinality of the agenda and the number of agents. Unanimity holds whenever $\odot$ is a function that satisfies minimality. Majority-preservation holds only for $\Delta^{d_H,\Sigma}$ and $\Delta^{d_T,\Sigma}$. The sequence-separability (S-s) holds for all aggregation function $\odot$ which are associative, however, as the example with the non-associative Π shows, there exist non-associative arithmetic aggregation functions that satisfy (S-s).

The literature of judgment aggregation, see for example (List and Polak, 2010; Pigozzi, 2006), discusses the anonymity and independence of irrelevant information for distance-based rules, but does not formally define or prove these properties since it is rather simple to show that they hold, and not hold respectively. The belief-merging operators are analyzed with re-

Property	Satisfied	Not satisfied
Unanimity	(d,Σ) (d,max) $(d,Gmax)$ (d,Π^*)	
Weak unanimity principle	$(d_D,\odot)$	(d_i,Σ) (d_i,max) $(d_i,Gmax)$ (d_i,Π^*) $d_i \in \{d_H, d_T\}$
Strong unanimity principle	$(d_D,\odot)$	
Majoritarian	(d,Σ) (d,Π^*)	(d,max) $(d,Gmax)$
Majority-preserving	(d_H,Σ) (d_T,Σ)	(d_D,Σ) (d,max) $(d,Gmax)$ (d,Π^*)
IR	$(d,\odot)$	
S-s	(d,Σ) (d,max) $(d,Gmax)$ (d,Π^*)	

Table 4.15: The summary of the properties which holds for the d and $\odot$ examples we introduced.

spect to unanimity and majoritarianism (Konieczny and Pino-Pérez, 1999). In addition to the aggregation functions and distances considered there, we also introduce the Π^*, which is majoritarian but not majority-preserving. The rest of the properties, to the best of our knowledge, have not been previously considered for judgment aggregation rules or functions.

Which properties should a rule aggregator satisfy? The structural properties are desirable in all the settings, as is the property of unanimity. In all consensual groups it is required that the adherence to majority properties are satisfied, as well as the unanimity principle at least in its weak version. The properties of insensitivity to reinforcement are particularly desirable in contexts in which the agents give judgments on the same issues several times, irrelevant of weather the group is consensual or hierarchical. A rule insensitive to reinforcement can save the agents from executing unnecessary aggregations. The properties of separability are particularly desirable for aggregators used by distributed consensual sub-groups. If a rule satisfies separability, then the smaller sub-group of agents can aggregate its judgments and send the result. The whole group may not need to aggregate the whole profile, but just consider these sub-results.

For group decision problems, ideally one would prefer resolute rules. However, from the impossibility results in judgment aggregation, see (List and Polak, 2010) for an overview, and social choice theory in general, we can conclude that only for restricted domains resolute aggregators can be constructed. If the domain cannot be restricted, and usually this is the case, then irresoluteness must be dealt with by tie-breaking mechanisms. If resoluteness is not feasible, then the rules should at least select a small number of collective judgment sequences as possible. For this reason, rules such as R_{IY}, R_{MR} and $\Delta^{d,max}$ are undesirable.

The rules we consider are in principle all desirable rules. As we can observe from Tables 4.14 and 4.15, the properties we considered are insufficient to fully distinguish among the rules. To this end more properties need to be developed. We made the initial efforts along this path in Section 4.10 where we discussed and defined five new rule aggregator properties. The family of interesting rule aggregator properties is still not large enough and the search for these properties is an open question in judgment aggregation theory.

In the third part of the thesis we consider instances of hierarchical and consensual groups in multi-agent systems and give judgment aggregation based models of decision-making for these groups. We pair rules from Chapters 2 and 3 with each of these decision-making problems using the properties we developed in this chapter.

Part III

Applying judgment aggregation rules in multiagent systems

5

Recognition-primed group decisions for hierarchical teams

Abstract. In this chapter we give an example of a hierarchical group decision problem in a multi-agent system context. When operating in uncertain environments, agents cannot rely on negotiation to reach agreements since the state of the world might change while they negotiate. We propose a model for reaching a group decision without negotiation. Our model lifts the Recognition-Primed Decision (RPD) model, constructed in an experimental psychology, from a single agent to an agent group. The lifting is executed by embedding judgment aggregation as a tool for amalgamating individual information. The RPD model models adaptive behavior. While it executes its actions, the group may adapt the decisions it acts upon in light of new information. We consider revision strategies for our group.

5.1 Introduction

Groups of agents need to be able to reach collectively binding decisions in order to coordinate and cooperate. We consider a hierarchical team of agents. In such a team, there exists one agent that is responsible for producing the group decision. To reach the group decision this agent needs to consider and combine the opinions of the rest of the team members. How can a hierarchical team reach collectively binding decisions in an uncertain environment?

According to traditional theory of decision-making, see for example (Peterson, 2009, Chapter 1), making decisions is driven by the concept of rationality associated with the decision-maker. A rational agent chooses, given his knowledge about the world, those options that are optimal in the sense that they maximize the agent's expected utility. Optimizing is difficult when the agents' resources are limited, as initially pointed out by (Simon, 1955, 1956). Furthermore, rationality is a concept associated with a decision-maker (an individual) and it is not simple to apply this concept to groups, see for instance the analysis of (Stirling and Nokleby, 2009).

People are not good rationalizers (Hardy-Vallèe, 2007, pg. ix), however groups of people are able to function successfully even when all adequate information is not available, when their goals are unclear and the procedures they have to follow are poorly defined, consider for instance firefighters and other emergency rescue teams. In life-threatening situations and dangerous environments it is desirable to replace human teams with artificial agents. Can artificial agents be endowed with such skills of decision making and adaptation that people possess? This question begets another question. How do people make collectively binding decisions under time pressure, in dynamic conditions and in uncertain environments? Can we use a model of human decision-making to build a procedure that can be used by artificial agents?

Using computational modeling, *i.e.*, multiagent based simulation of group-decision making theories built and studied in experimental psychology, is used to validate and analyze human decision-making models (Ilgen et al., 2005; Hulin and Ilgen, 2000). The less explored direction is the use of experimental psychology models to build decision-making or agreement reaching procedures for artificial agents. One reason for this might be found in the non-simplicity and high non-determinism of the experimental models.

Compared to studies of consensual groups, the hierarchical team decision-making is far less studied in experimental psychology and social sciences (Humphrey et al., 2002). A summary of theories on how hierarchical team-decision making is done, or should be done, is given in (Humphrey et al., 2002). A well known model is the *multi-level theory of team decision-making* of (Hollenbeck et al., 1995). This theory however is rather intricate and it would be difficult to translate into a group-decision making model for artificial agents.

How firefighter commanders make decisions under extreme time pressure was studied by Klein et al. (2010). They found that, when a commander has prior experience with a problem, which is usually the case, he acts according to the *recognition-primed decision* (RPD) model, summarized on Figure 5.1.

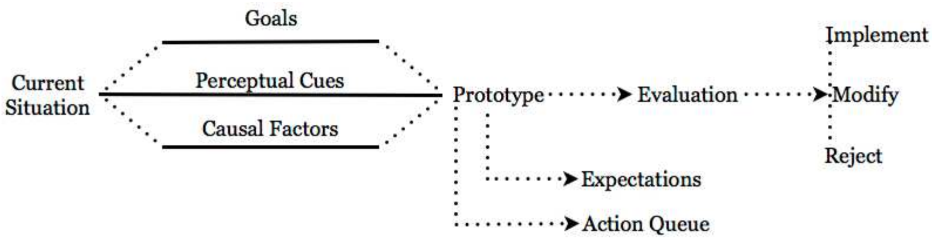


Figure 5.1: The recognition-primed decision model (Klein et al., 2010, pg. 203).

According to the RPD model, a commander tasked with a problem first assess the current situation and then matches the current situation to a prototypical solution based on similarity of goals, perceptual clues, causal factors and information about them.

As a running example we consider the overpass rescue example given in (Klein, 1999, pg. 18).

Example 5.1.1. *The overpass rescue(Klein, 1999, pg. 18).*

> *"A lieutenant is called out to rescue a woman who either fell or jumped off a highway overpass. She is drunk or on drugs and is probably trying to kill herself. Instead of falling to her death, she lands on the metal supports of a highway sign and is dangling there when the rescue team arrives.*
>
> *The lieutenant recognizes the danger of the situation. The woman is semi-conscious and lying bent over one of the metal struts. At any moment, she could fall to her death on the pavement below. If he orders any of his team out to help*

her, they will be endangered because there is no way to get a good brace against the struts, so he issues an order not to climb out to secure her.

Two of his crew ignore his order and climb out away. One holds onto her shoulders and the other to her legs.

A hook-and-ladder truck arrives. The lieutenant doesn't need their help in making the rescue, so tells them to drive down to the highway below and block traffic in case the woman does fall. He does not want to chance that the young woman will fall on a moving car.

Now the question is how to pull the woman to safety.

First, the lieutenant considers using a rescue harness, the standard way of raising victims. It snaps onto a person's shoulders and thighs. In imagining its use, he realizes that it requires the person to be in a sitting position or face up. He thinks about how they would shift her to sit up and realizes that she might slide off the support.

Second, he considers attaching the rescue harness from the back. However, he imagines that by lifting the woman, they would create a large pressure on her back, almost bending her double. He does not want to risk hurting her.

Third, the lieutenant considers using a rescue strap-another way to secure victims, but making use of a strap rather than a snap-on harness. However, it creates the same problems as the rescue harness, requiring that she be sitting up or that it be attached from behind. He rejects that too.

Now he comes up with a novel idea: using a ladder belt - a strong belt that firefighters buckle on over their coats when they climb up ladders to rescue people. When they get to the top, they can snap an attachment on the belt to the top rung of the ladder. If they lose their footing during the rescue, they are still attached to the ladder so they won't plunge to their death.

The lieutenant's idea is to get a ladder belt, slide it under the woman, buckle it from behind (it only needs one buckle), tie a rope to the snap, and lift her up to the overpass. He thinks it through again and likes the idea, so he orders one of his crew to fetch the ladder belt and a rope, and they tie it onto her.

In the meantime, the hook-and-ladder truck has moved to the highway below the overpass, and the truck's crew members raise the ladder. The firefighter on the platform at the top of the ladder is directly under the woman shouting, 'I've got her. I've got her.' The lieutenant ignores him and orders his men to lift her up.

At this time, he makes an unwanted discovery: ladder belts are built for sturdy firefighters, to be worn over their coats. This is a slender woman wearing a thin sweater. In addition, she is essentially unconscious. Then they lift her up, they realize the problem. As the lieutenant put it, "She slithered through the belt like a slippery strand of spaghetti."

Fortunately, the hook-and-ladder man is right below her. He catches her and makes the rescue. There is a happy ending.

Now the lieutenant and his crew go back to their station to figure out what had gone wrong. They try the rescue harness and find out that the lieutenant's instincts were right: neither is usable.

Eventually they discover how they should have made the rescue. They should have used the rope they had tied to the ladder belt. They could have tied it to the woman and lifter her up. With all the technology available to them, they had *forgotten that you could use a rope to pull someone."*

The recognition-primed decision model captures the behavior of the lieutenant in Example 5.1.1. After arriving to the scene, the lieutenant first assesses the situation. He observes that the problem is "*. . . a woman who either fell or jumped off a highway overpass. She is drunk or on drugs and is probably trying to kill herself. Instead of falling to her death, she lands on the metal supports of a highway sign and is dangling there . . .* ", and "*. . . The woman is semiconscious and lying bent over one of the metal struts. At any moment, she could fall to her death on the pavement below.*"

The lieutenant matches the current situation to a prototype based on similarity of goals, perceptual cues, causal factors and information about them. Namely, he is "*called out to rescue a woman*", "*he does not want to chance that the young woman will fall on a moving car.*" The matched prototype generates *expectancies* and a set of options for a course of action. The expectancies are a mean of confirming that the selected prototype is adequate. The options are generated sequentially, with the most typical option being generated first and other options only being generated if the previous one is rejected. In the overpass rescue example, the first generated option by the lieutenant is to use a rescue harness in a standard manner. The second generated option is to attach the rescue harness to the victim from the back. Once a course of action is generated, the commander proceeds to evaluate it for plausibility and implements it, modifies it, or rejects it. In the overpass rescue example, the lieutenant evaluated the use of a rescue harness and rejected it because "*it requires the person to be in a sitting position or face up* and that is not the case. If an option is rejected, the next most available, representative, and similar one is selected for evaluation.

The RPD model is a relatively simple model and it was developed to describe the behavior of human resource limited teams. Decisions in conditions like the ones in which the firefighters operate must be made fast. Furthermore, the team that makes them must be able to adapt easily. This is precise the quality of group decision-making that we search for and that cannot be accomplished by traditional optimization based decision-making. However, there are two problems in using the RPD model to build a group decision-making procedure for artificial agents. The first problem is that the commander in the RPD model uses his experience and associations to match a possible solution and cues that verify it as an adequate solution. This task is more difficult to perform by an artificial agent than people since people can use a small number of cases and associations to find solutions for a given problem. The second problem is that although firefighters operate as teams, the RPD is a model of a single agent. Consequently, a team decision-making procedure based on the RPD model will be only applicable to groups that solve problems on which they have prior experience.

We want to replace the agents on the ground with artificial agents like robots and drones, but we can use a person, let us call this person *an initiator*, to perform the task of the commander. This way we circumvent the first problem. Since the initiator is no longer on the scene of the event, he or she would not be able to assess the situation and verify the expectancies. We need to lift the RPD model from a single agent to a hierarchical group model. In this chapter we address this lifting problem.

We consider a mixed human-robot team in which there is one human, called *initiator*, which has a role similar to the firefighter commander's role. The rest of the agents are artificial agents, called *executors*. Unlike the commander, the initiator is not on the ground and has to fully rely on the executors for the following processes:

1. situation assessment,
2. verifying expectancies, and
3. evaluating potential courses of action.

The challenge in raising the recognition-primed decision making model to the team level lies in raising these processes to the team level. An adequate collective decision process, or processes, needs to be specified to accomplish this task.

The recognition-primed decision approach is very fast; Klein estimated that the fire-ground commanders make around 80 percent of their decisions in less than a minute (Klein, 1999, pg.4). Inevitably, the decision-making can be expected to take longer when there are opinions from various sources to be merged. Collective decisions in multiagent systems can be reached by argumentation supported negotiation (Rahwan et al., 2003) and social choice (Chevaleyre et al., 2007). Negotiation requires several rounds of exchanges (of arguments) between the agents before a decision is reached. It can be used in software agents, but not yet for robotics. Unlike software agents, embodied agents need to assess the environment, process their sensor input and form an opinion. A robot that has to form the opinion, not just pull it out of his knowledge base, would find itself in “no time to think”. The agreements reached by embodied agents under time pressure must be done with as little information exchange as possible. The executor needs to be able to get all, or at least most, of the information from the agents at once and deduce the courses of action from it.

Social choice methods such as voting (Nurmi, 2010) and judgment aggregation (List and Puppe, 2009) require only one round of exchange of information. The initiator can apply a social choice rule to aggregate the executors’ opinions. In the case of the situational assessment, the agents need to express a judgment whether the cue is present or not. However, not all executors would be in a position to make a judgment on all cues. The opinions on the cues entail the opinions on the solution that can be applied, but the judgment made on some cues may logically constrain the judgments that can be made on the others. Therefore the initiator should use judgment aggregation rules, in particular weighted rules for ternary judgments like the ones we developed in Chapter 3.

This chapter is structured as follows. In Section 5.2 we propose a group decision-making model based on the recognition-primed decision model. In Section 5.3 we focus on the problem of reaching collective decisions by judgment aggregation within our model. In Section 5.4 we revisit the overpass rescue example and show how our model can be applied to it. In Section 5.5 we study the problem of revising the emerging states with new information. In Section 5.6 we present our conclusions and discuss possible generalizations of the proposed model to teams with no initiator.

5.2 A conceptual model of reaching recognition-primed group decisions

We construct a conceptual model of recognition-primed group decisions for the mixed human-robots team. The model works under the assumption that all agents are able to communicate with each other. We begin by describing the possible roles in the team and the presumed capabilities of each role.

The model we present here is a prescriptive model for a team recognition-primed reasoning for collaborative problem solving in uncertain environments. The recognition-primed decision-making model of (Klein et al., 2010) is a prescriptive model of a single agent decision-making in uncertain environments.

5.2.1 The team

The initiator

We define the initiator as an agent who is able to use his experience in matching a given problem to a pair of sets: set of goals and a set of corresponding cues for each goal. The cues, corresponding to one matched goal, identify when this goal is good enough to be adopted in response to the given problem. For example, given the problem of rescuing the unconscious woman, the goal to lift her to safety using a rescue harness should be pursued if the woman is facing up or is in a sitting position. It is not sufficient to only enlist the cues for a matched goal. The relational structure between the cues and the goal needs to be specified as well.

The initiator is able to match a problem with the triple $\langle goal, cues, relational\ structure\rangle$. In the remainder of this chapter we will mean both cues and the relational structure when we speak of a set of cues.

Once a goal is matched, and verified as good enough by evaluating the cues, the initiator constructs a plan for that goal. We assume that the initiator is an agent able to generate plans.

The executors

An executor is an agent who is able to generate an opinion for a given cue based on his own knowledge, beliefs and percepts. He is able to evaluate the role assigned to him by the initiator's plan and identifies the constraints that would inhibit the successful execution of the tasks assigned to him. The executors are able to pass messages between each other in order to successfully execute a plan. For instance, during the (hypothetical) execution of the plan for rescuing the woman by a rescue harness the agent that straps the harness needs to signal the agent that lifts the woman that he can start lifting.

5.2.2 The process

The left side of Figure 5.2 represents the recognition-primed group decisions (RPgD) model for the initiator. The right side depicts the model for the process for the executor agent.

The process begins once the initiator recognizes a problem, or is tasked with one. He contacts the executor agents who are already on site or on stand-by, to determine who is available to participate. If the initiator finds sufficient agents, he proceeds to establish a course of action. He first identifies the team goal(s) by using his experience to find the closest match of the problem at hand with a goal. The initiator also matches the relational structure and corresponding cues to the proposed goal.

Situational assessment

The next step is the situational assessment with the purpose of goal verification. This step comes only after the executors arrive at the problem site, if they are not already there. The initiator assesses the situation by eliciting the executors' opinions on whether the cues are true or false, present or not. The cues are thus treated as propositions to which an executor assigns true or false. The executor can also abstain from assigning a value. Based on the reported cue value assignments, the initiator assigns corresponding cue true/false values which he uses to establish whether the corresponding goal is good enough or not and adopted or not correspondingly. We propose that the initiator uses a judgment aggregation rule to aggregate the reported information. If negotiation were to be used to assess the situation, then the executors would first reach the agreement of each cue value between them and then report this agreement to the initiator.

The initiator verifies the adequacy of the matched goal, based on the corresponding cue values obtained by aggregation. He can request only the opinions on the cues to be aggregated, or he can consider also the aggregation of the individual conclusions on whether the goal should be adopted.

Example 5.2.1. *Consider the problem of to pulling the woman, from the overpass rescue example, to safety. Assume that there are five executors* $\{E_1, E_2, E_3, E_4, E_5\}$*. The initiator first considers the goal*

s_2*: use a rescue harness.*

Based on his experience, the initiator deems s_2 *a satisfactory solution if and only if at least one of the following cues are the case:*

c_2*: the victim is in a sitting position,*

c_3*: the victim is in a face-up position,*

c_4*: the victim can safely be shifted in a sitting up position or in a face-up position,*

c_5*: the harness can be attached from the back without hurting the victim.*

The initiator further specifies that the cues are subject to the constraint $(c_2 \vee c_3) \rightarrow c_4$*. The constraints encodes the "obvious" information that if a victim is already sitting or facing up, then* c_4 *is trivially the case. The judgments of the executors, regarding the cues, and the individually entailed judgments regarding* s_2 *are given in Table 5.2.1. The "?" denotes the case in which the agent has not provided a judgment.*

Agents	c_2	c_3	c_4	c_5	s_2
$\{E_1, E_2\}$	no	no	no	no	no
$\{E_3, E_4\}$	no	no	?	?	?
$\{E_5\}$	no	no	no	?	?

Table 5.1: Contributed judgments regarding cues on s_2.

In Example 5.2.1 the individual judgments regarding the goal adequacy can be deduced. An opinion of a goal given explicitly by an executor carries additional information. Assume, for example, that in the situational assessment for goal s_2, agents $\{E_3, E_4\}$ give an explicit opinion "yes" on s_2. The relation structure is still verified. However, additional information is conveyed. Namely these agents are of the opinion that at least one of c_4 and c_5 must be the

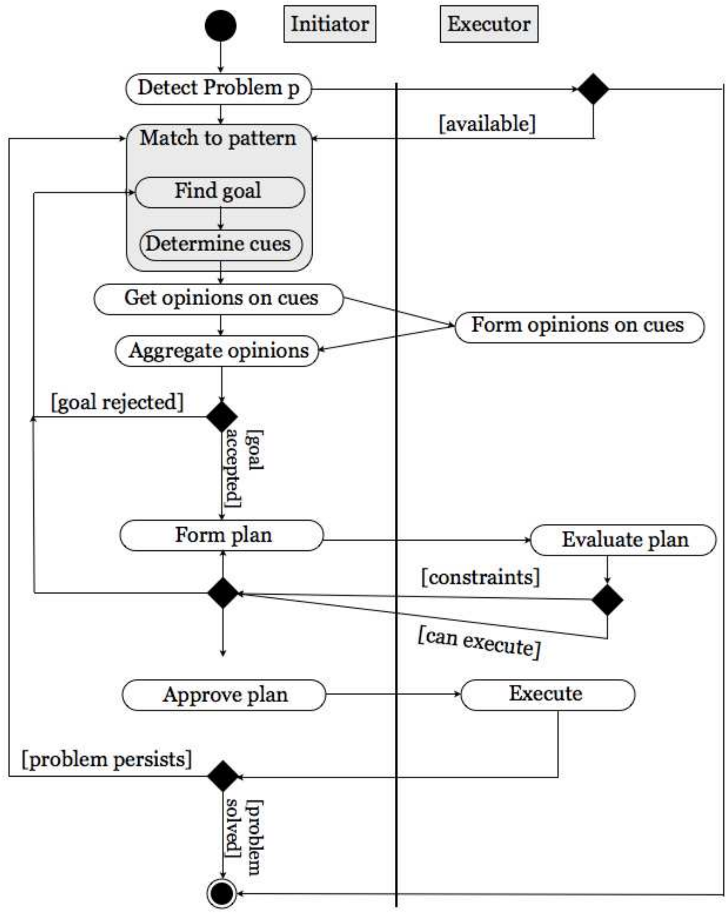


Figure 5.2: The process of making recognition-primed group decisions.

case, but they are unsure which one. If the initiator wants to include these "unsupported by cues" opinions, then he should elicit opinions on s_2 explicitly.

If a goal is not adopted, then the executor generates another goal (and corresponding cues) and elicits the opinions of the agents for this set. If the new goal shares cues with a goal previously considered, then there is no need to elicit the judgments of the executors on these cues anew. The initiator keeps generating goals and considering them one by one until a goal is adopted, as long as the problem persists. If the initiator runs out of ideas for a possible solution, then the project can be abandoned, or the initiators can remain on scene expecting for the situation to change. In the firefighter examples collected by Klein, the commander always has an all-contingency solution. Thus, for example if a fire cannot be extinguished, and no rescue is deemed possible, then the firefighters take actions to ensure that the fire can be left to burn itself out safely. Similar all-contingency solutions can be designed for each problem domain.

Plan evaluation

Once a goal is deemed adequate, the initiator proceeds to form a plan. The initiator has information about the scope of capabilities of an executor, but he does not know the exact position or state of every executor at a given point in time. Consequently, the initiator makes a tentative plan based on the information of executor capability he has. He proposes this tentative plan to each executor.

An executor has two options. One option is to acknowledge the plan, and the role assigned to him in that plan. By this action he commits himself to executing his assignments. The second option is to object to the plan. An executor objects by informing the initiator of the constraints due to which the proposed plan cannot be executed. We observe that the mental simulation, which is done by the commander in the model of Klein, is, in our model, externalized from the initiator to the executors.

The initiator considers all constrains given as an objection to a plan in order to devise a new plan. A plan is approved if and only if there are no objections from any of the executors. If the constraints are such that no plan can be devised for the goal in question, the initiator attempts to matches a new goal to the problem.

During plan evaluation, an agent evaluates the portions of the plan, which he is expected to execute. The plan may contain both individual and joint actions that are to be executed. When the action is individual, no conflict of constraints can arise, since the agent who is intended to execute the action is taken to have a veto on evaluating the plan for that point.

When the action is joint, all agents involved need to approve the tentative plan, for the plan to adopt. An agent is allowed to objects on an action that is not executed by him. There are two types of objections that can be raised. The first objection is due to a conflict in the execution of the assigned actions. The second is an objection to the abilities of another agent. In the case of the latter objection type, the initiator can approach the objection acting on *a cautious mode* or on a *a brave mode*. In a cautious mode he will never approve a plan as long as there is an objection to it. In a brave mode, the initiator can consider the agent who executes the action to be the ultimate authority regarding his abilities, and disregard these types of objections.

Once a plan is approved, the executors proceed with the tasks that are assigned to them. After the plan is executed, if the problem still persists, then the initiator attempts to match the problem with another goal, and verify it based on its corresponding cues. A decision to adopt or refute a goal, values assigned to cues, or a plan, can be reconsidered as long as the problem persists.

5.2.3 Group recognition-primed decision-making and satisficing

Simon (1955, 1956) addressed the question of how a resource bounded agent makes decisions. He argued that a resource-bounded agent should not maximize expected utilities, but select the first option that is good enough. An option is good enough when a sufficient number of indicative criteria are satisfied. This process of selecting the first good enough option he called “satisficing”. In contrast to satisficing, traditional decision-making, see for example (Peterson, 2009, Chapter 1), is the process of first enumerating all possible solutions for a problem and then selecting from them that solution which is optimal in the sense that it maximizes the expected utility of the decision-maker. Since its inception, satisficing has gathered considerable attention and many variations of this concept have been developed (Radner, 1975; Matsuda and Takatsu, 1979; Shinzo and Takatsu, 1980; Wierzbicki, 1982; Haller, 1985; Kaufman, 1990; Brown, 1990; J. and L., 1990; Zilberstein, 1998; Greiner et al., 2006; Güth, 2010). There is not much study on satisficing for groups, an exception being (Stirling and Nokleby, 2009). It is not our intention to analyze all these models and approaches, but highlight the advantages of satisficing, particularly for decision-making in groups, and the satisficing aspects of the recognition-primed model.

A satisficing solution is not necessarily an optimal solution in the utilitarian sense. *E.g.*, there is no evidence that suggests that the solution to use the rope to lift the victim into safety, in the overpass rescue example, is necessarily the fastest, cheapest, or safest solution possible. The lieutenant did not exhaust all the options for saving the unconscious victim; she could have been pushed on an inflatable trampoline, lifted by a helicopter, etc. The advantages of pursuing optimal decisions are evident, however in situations in which there is no time to generate all options and evaluate them, satisficing is a better strategy. Because a course of action is determined fast, satisficing allows the team to be more adaptive to changes in its environment. Simon proposed the concept of satisficing but he did not propose a formal model.

Another argument for using satisficing instead of optimizing in some multi-agent systems settings is that of problems that arise with the concept of rationality. Decision-making is driven by the concept of rationality associated with the decision-maker. Rationality is a property of an individual, regardless of whether that individual is taken to be one agent or one team of agents. In the case of group decision-making, it might be problematic to identify how to apply the concept of rationality (Stirling and Nokleby, 2009). For the concept of rationality to be successfully applied to the team members as individuals, the agents must be assumed to be perfectly competitive. As Arrow (Arrow, 1986, pg. S387) observed:

> “rationality in application is not merely a property of the individual. Its useful and powerful implications derive from the conjunction of individual rationality and the other basic concepts of neoclassical theory - equilibrium, competition, and completeness of markets. [...] we need not merely pure, but perfect competition before the rationality hypotheses has their full power. [...] When these assumptions fail, the very concept of rationality becomes threatened, because perceptions of others and, in particular, of their rationality become part of one’s own rationality.”

For the concept of rationality to be applied to the team as a unit, the agents in the team must be in perfect cooperation, in the sense that none of them has goals that are not goals of the team. Observe, for instance, that the safest way of rescue for the victim is not necessarily the safest way of rescuing the victim for the firefighters. While the team is cooperative, the members of the team must maintain some level of self-interest when it comes to ensuring their own safety.

Satisficing, as a concept, can be seen as predominantly associated with the course of action itself, rather than with the agent who selects the course of action. The solution for the problem of rescuing the unconscious woman is the one that meets the minimal conditions to be adopted: it gets the job done and it can be done by the firefighters. The rescue solution that is optimal needs to maximize both the utility of the team, which can be seen as predominantly cooperative, and the utility of the unconscious woman, whose utility can be seen as competitive with respect to the firefighter. Applying satisficing to groups, regardless of whether they are cooperative or competitive, is not more problematic than applying it to individuals, as long as the groups have a way of determining what are the sufficient conditions and whether they are satisfied. In a hierarchical group, such as the firefighting team, the initiator makes this decision.

5.2.4 Team adaptation and recognition-primed reasoning

The main characteristic of the firefighters studied by Klein is that they constitute highly adaptive teams. Being adaptive is a necessary property of teams that operate in an uncertain environment. Burke et al. (2006) define team adaptation as a change in team performance, in response to salient cues, that leads to a functional outcome for the team. It is further specified that "team adaptation is manifested in the innovation of new or modification of existing structures, capacities, and/or behavioral or cognitive goal-directed actions" (Burke et al., 2006, pg.1190).

The adaptive cycle of the team adaptation model presented in (Burke et al., 2006) is characterized by four core constructs:

1. situation assessment;
2. plan formulation;
3. plan execution, via adaptive interaction processes;
4. team learning.

The adaptive cycle is further characterized by *emerging cognitive states*, such as shared mental models, "which serve as both proximal outcomes and inputs to this cycle", (Burke et al., 2006, pg. 1192).

The recognition-primed agreement model we propose verifies the team adaptation model of Burke et al. (2006). The emerging states in the case of our model are the agreements regarding goals, value of cues and the adopted/refuted plans. The verification of expectancies and the valuation of plans are the way in which the emerging states are reconsidered. In the context of our recognition-primed model, the actions of the team are based on the emergent states and the team adaptation is a result of the adaptation, or reconsideration, of the emergent states. In Section 5.5 we discus the reconsideration of the emerging states.

The process we do not explicitly consider in our model is learning, since we focus on giving a conceptual model of reasoning and not of learning. However, we can observe that from the aspect of improving team performance, learning is an important process both for the initiator and the executors. The initiator can improve his accuracy in matching a problem with a goal and cues, while the initiators can learn to improve their plan evaluation and cue observation accuracy.

The recognition-primed decision model of Klein does not explicitly include the process of learning either, as it can be observed even on the more detailed depiction of the model given on Figure 5.3.

However, in the original overpass rescue example, after the problem is solved, *i.e.*, the woman is rescued, the lieutenant analyzes the situation anew to determine that the best course of ac-

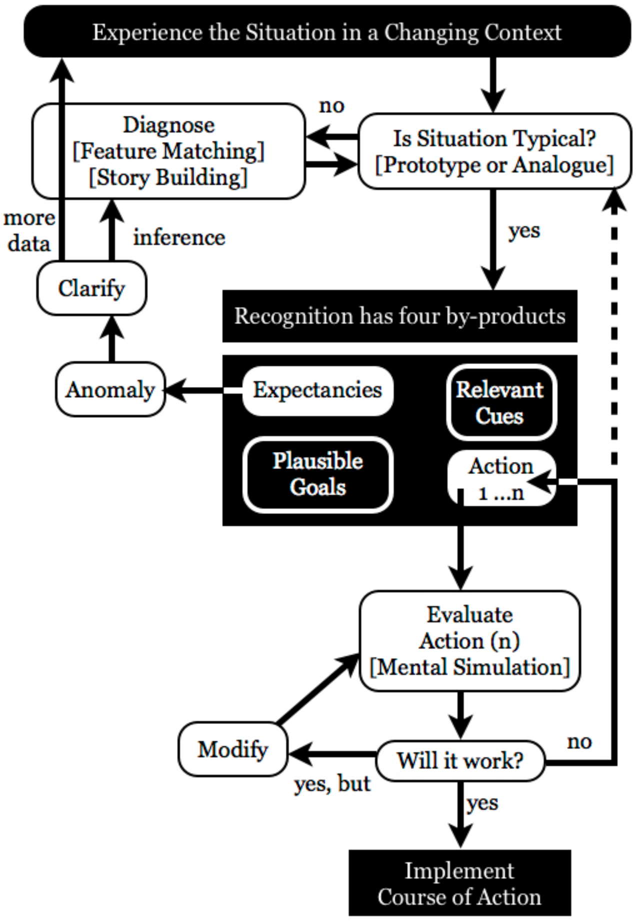


Figure 5.3: The recognition-primed decision model as given in (Klein, 1999, pg. 27).

tion was to use a rope to lift the victim. This is when the lieutenant *learns* from the experience by accommodating this new experience with the old ones.

5.3 Collective decision-making problems in the RPgD model

In this section we consider the collective decision-making problems that occur in the scope of the group recognition-primed decision-making model. The result of the group-decision making are the emerging states which are the decisions on whether to adopt a given plan and the decision regarding the situation assessment.

5.3.1 Emerging states and reaching agreement on a plan

In our model the initiator depends on the perceptions and opinions of the executors to assess the situation, verify the expectancies and evaluate a tentative plan. Each executor can have different knowledge, beliefs and perceptions of the world, which give rise to the possibility that the executors will give different values to different cues, different opinions on whether a goal is to be pursued and different views regarding whether a plan is executable. The initiator needs to "reconcile" the different opinions to be able to coordinate the activity of the team. The emerging states, *i.e.*, , the presently established collective cue values, adopted goals and approved plans, are obtained as an end product of this "reconciliation" process.

The information requested and submitted regarding the cues and the goals is of different nature than the information exchange regarding a plan. The cues, and the goal they correspond to, are binary questions to which an executor answers with "yes", when he thinks the cue is present, "no", when he thinks that a cue is not present, or with abstaining from giving a "yes" or "no". The tentative plan evaluation is an information request to which two types of reply are possible: either an approval of the plan or a constraint which indicates the plan's unsuitability.

All agents who are tasked with giving a cue, or opinion on a goal, are expected to produce an answer, and "yes" and "no" have different meanings from abstaining to reply. Even if the agents are divided regarding whether a goal should be adopted, the initiator may conclude to adopt the goal. In contrast, the plan evaluation operates according to *qui tacet consentire videtur* [1]. The constraints from all the agents are taken into account when the initiator forms the next plan. As long as there is at least one person who objects, a plan cannot be approved. Due to the latter, the decision to approve a plan is rather simple. A decision on a situation assessment poses more of a challenge. In the next section we focus on the problem of reaching these group decisions.

5.3.2 Situation assessment as a judgment aggregation problem

How can the initiator form the decision on whether a goal is adequate or not? The initiator can first inform the agents of the relational structure that shows how an opinion on a goal can be deducted and then ask the agents to deduct their own judgments on the goal instead of assessing the situation by eliciting judgments on the cues. The judgments on the goals can

[1] A legal expression used to state the convention that in negotiations, the one who has noting to say is taken to be in agreement with what is proposed.

then be pooled to determine whether a goal is to be pursued or not. However, this approach is not an option for adaptable teams. If agreements on the cues are not reached, or known, then it is difficult to update the goal when the state of the world changes.

When the world changes, it may not be obvious how that affects this goal adequacy. The cues act as criteria for evaluating the adequacy of the proposed goal for the problem at hand. When the decisions on cues are reached, the changes in the state of the world can be deemed relevant when the decisions on the cues are inconsistent with them. Not only do the cues show when a goal should be reconsidered, but also how to be reconsidered. This is why it is important to reach decisions on the cues as well as on the goals and why we use judgment aggregation to determine them.

A judgment aggregation problem is specified by an agenda, a set of constraints and a finite set of agent names N. The representation of the situation assessment agreement problem in judgment aggregation is straightforward.

The judgment aggregation problem is represented using a logic $\mathcal{L}$ and an entailment relation for that logic $\models \subseteq \mathcal{L} \times \mathcal{L}$. The agenda and the constraints are sets of well formed formulas from this logic.

Since the goals are considered for adequacy sequentially, the agenda will contain one goal and arbitrarily many cues. An exception is the case when the goals generated are concurrent and non-conflicting, in which case they will be generated at the same time and considered both in the same agenda.

For example, the agreement problem for goal s_2 can be represented in propositional logic with agenda $\mathcal{A} = \{c_2, c_3, c_4, c_5, s_2\}$ where the cues and goal are represented by propositions, and $\mathcal{R} = \{(c_2 \vee c_3 \vee c_4 \vee c_5) \rightarrow s_2\}$. We need a ternary logic to represent the judgments. The profile π of individual judgments for $\mathcal{A}$, according to Example 5.2.1, are

$$\begin{array}{r} \\ E_1, E_2 \\ E_3, E_4 \\ E_5 \end{array} \begin{array}{c} \begin{array}{ccccc} c_2 & c_3 & c_4 & c_5 & s_2 \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \end{array}.$$

Which ternary logic semantics should we use for representing aggregation problems in the RPgD model? Many ternary logics have been presented in the literature, see for example (Urquhart, 2001) for an overview of the ones considered basic. In a ternary logic, in addition to the values assigned to "true" and "false" there is a third value between them. The difference between the different ternary logics is in the mathematical and philosophical semantics attached to the intermediately value. In representing judgments on cues, we use the third value to represent the following cases:

- when an agent abstains from making a judgment on a cue,
- when an agent abstains from sending an opinion whether a goal should be adopted or such opinion is not deducible from his judgments on the related cues,
- when a collective judgment on a cue, or goal adoption cannot be determined by the aggregation rule used.

We consider the basic ternary logics and discuss their adequacy for use as a representational language for judgment aggregation for situation awareness. The logics we consider are:

- the logic of Post (1921),
- the logic of Bochvar (1938),
- the logic of Łukasiewicz (1920)[2], and
- the logic of Kleene (1938).

Post (1921) was one of the first to introduce a many-valued logic. In his t-valued system, the propositions are assigned values from $T = \{0,\ldots,t-1\}$. The lowest value corresponds to the lowest degree of truth and the highest to the highest degree of truth. Post constructed his logic in a purely mathematical manner and did not attribute any philosophical analysis to any of the intermediate values. The semantics of the propositional logic operators is given as:

$$v(\neg\psi) = v(\psi) + 1(\text{mod t})$$

$$v(\varphi \wedge \psi) = min(v(\varphi), v(\psi)).$$

For the case of $t = 3$ we obtain a ternary logic in which, 0 denotes “false”, 2 denotes “true” and 1 is assigned to the intermediary value.

Post’s is not a good logic for representing judgments in the situation assessment context. Observe that a negated intermediate proposition is assigned the value “true” and the negation of a “false” proposition is assigned the value “intermediate”. Consider the Example 5.3.1.

Example 5.3.1. *An executor is asked to give judgments on the propositions:*

a_1 *the victim is conscious,*

a_2 *the victim is in a safe location,*

and then, based on these judgments, using $\mathcal{R} = \{(\neg a_1 \wedge a_2) \leftrightarrow s\}$ *the initiator deduces this executor’s opinion on*

s *send for a ladder-truck.*

Assume that the executor from Example 5.3.1 reported the judgment sequence $A = \langle 0,2,1\rangle$, namely that he finds the victim unconscious and in a stable location. According to the semantics of Post, the value assigned to s is 1, meaning that the executor recommends abstaining from a course of action even though he did not abstain on any of the judgments that determine if a course of action is to be adopted.

Recall that a judgment sequence is consistent in propositional logic, if for the corresponding $\hat{A}$, it holds $\hat{A} \cup \mathcal{R} \not\models \perp$. We can extend the definition of consistency for a classical ternary logic entailment operator $\models_3$ (Cadoli and Schaerf, 1996), straightforwardly. Observe that if an agent i assigned a value 1 to an issue a, then $a \notin \hat{A}_i$ and $\neg a \notin \hat{A}_i$.

In ternary logic, we can have $\hat{A} \cup \mathcal{R} \models_3 \perp$ being true, false or undecided. A judgment sequence is inconsistent if and only if $\hat{A} \cup \mathcal{R} \models_3 \perp$ is false. This means that, if such a $\models_3$ is used, the executor could have submitted the judgment set $A = \langle 0,2,0\rangle$ as well, since it would

[2] See (Borkowski, 1970) for an English translation and (Urquhart, 2001) for a summary.

be consistent. This possibility makes it difficult for the executor to determine what is the decision regarding the goal.

The negation semantics according to Post is incompatible with the intuitive meaning of abstention, and as consequence, all ternary logics that have this negation semantic are unsuitable for the representation of judgment aggregation problems.

Bochvar (1938) defended the stance that the intuitive meaning assigned to the third, or intermediate, value is "meaningless".

According to the Bochvar logic semantics, a value "meaningless" is assigned to every formula that contains a proposition that is assigned the value "meaningless". Consequently, when the executors are not asked to explicitly state their opinion on the goal, Bochvar logic cannot be used. Consider again Example 5.3.1, but now assume that the executor reported the sequence $A = \langle 0, \frac{1}{2} \rangle$. According to the semantics of the Bochvar logic, the sequence A is consistent regardless of which value is assigned to s, since the value assigned to $\neg a_1 \wedge a_2$ is $\frac{1}{2}$. For any value assigned to s the value assigned to $(\neg a_1 \wedge a_2) \leftrightarrow s$ is $\frac{1}{2}$. When opinions about the goal are not directly elicited by the executors, the agents can declare whichever value for the goal.

Bochvar logic may still be used to represent the individual judgment sets, in the case when the initiator does not care about the individual values assigned to the goal and the cues are logically not related between themselves, as is the case in Example 5.3.1. However, using Bochvar logic to represent the collective judgment set is ill-advised since, as soon as there is an abstention in the collective values of the cues, the initiator will not be able to determine whether a goal should be adopted.

Łukasiewicz (1920) proposed his ternary logic independently from Post and unlike Post took the philosophical approach to developing it. According to the semantics he proposed, the formulas are assigned values from $T = \{0, \frac{1}{2}, 1\}$ with $v(\top) = 1$, $v(\bot) = 0$ and the intermediate value $\frac{1}{2}$ assigned to propositions whose truth state is "possible" or "to be determined later".

Unlike in the logic of Post, here the negation of the intermediate value is the intermediate value, while the negations of $\top$ and $\bot$ are as in classical logic. This is adequate for representing "true"/"false" judgments and abstentions. Unlike the logic of Bochvar, a value of an expression that contains a proposition assigned a value $\frac{1}{2}$ is not necessarily $\frac{1}{2}$. *E.g.,* if that the collective judgments on the cues for Example 5.3.1 are $v(a_1) = 0$ and $v(a_2) = \frac{1}{2}$, then the consistent collective judgment set is $A = \langle 0, \frac{1}{2}, 0 \rangle$. An initiator can decide on a adopting a goal in some cases, even if the collective judgment on some of the cues is $\frac{1}{2}$.

Using the ternary logic of Łukasiewicz is adequate for representing abstentions that occur when an agent is undecided regarding a cue or a goal. Consider for example the case when a robot has to make a judgment on whether an object is red or not. To do so he has to sample several readings from his wave length sensor and make a judgment "yes" if the average value of the readings is greater than 620*nm*. It can happen that his sensors give contradictory readings in the samples taken and as a consequence then the robot cannot set a judgment (without making further analysis). However, if after some period he is asked again for a judgment on the same cue, the robot might be able to produce a judgment.

The ternary logic proposed by (Kleene, 1938) is another good candidate for representing judgments in the situation assessment context. Kleene assignees the meaning of "unknown" to the intermediate value. The difference between the semantic of the Kleene logic and that of Łukasiewicz is in the interpretation of the implication $\rightarrow$: while Łukasiewicz deems $\frac{1}{2} \rightarrow \frac{1}{2}$ to be an expression that is evaluated to 1, according to Kleene, the same expression is evaluated to $\frac{1}{2}$.

Using the ternary logic of Kleene is adequate for representing abstentions that occur when an agent has no means to determine the judgment regarding a cue or a goal. For instance, a robot that has no microphone cannot determine the value for a cue "sound is coming from the room" and will report a value $\frac{1}{2}$. If after some period the need for a judgment on the same cue arises, provided he repaired his sensors, the robot will be able to produce a judgment.

It is not strictly necessary that all the agents use the same ternary logic semantics, but the $\models_3$ for the collective judgment sequence must be set so that the set of all consistent judgment sequences $\mathbb{A}(\mathcal{A},\mathcal{R},\models_3)$ can be determined. However, using different semantics makes the reasoning process more complex since the agents have to report the semantics they use. Furthermore, for different cues the agents may have different reasons for abstaining.

In many scenarios it can be expected that an executor is competent with respect to some cues and not so competent with respect to others. For instance, a robot can be better able to determine the position of the victim if he is closer to the victim. The added accuracy can be due to some particular expertise of the executor. A robot equipped with an infrared vision can be more precise in estimating whether an immobile victim is dead or alive that a robot making the same opinion based on movement recognition. A weight can be assigned by an initiator, or it can be provided by the executors themselves. In addition to the profile of judgments, a profile of weights is also going to be available to the initiator.

5.3.3 Rules for aggregating judgments for situation assessment

Which judgment aggregation rules should the initiator use to aggregate the individual judgments on the cues and the goal?

The situational assessment aggregation problem is such that a weighted ternary judgment aggregation rule is necessary, such as the ones we developed in Chapter 3. Since different goals, cues and relational structure can be specified, the judgment aggregation rule needs to be able to handle all agendas without constraints. Also, the agents can submit any combination of judgments, thus the rule needs to satisfy the universal domain property. Coordinating the input of the agents towards certain types of profiles requires additional communication thus slowing down the agreement process.

Based on these requirements, the initiator needs to use the rules $\Delta^{d,\odot}$ specified by Definition 43 for this X as a constraint for the co-domain. The best choice for d and $\odot$ are d_T and $\sum$ correspondingly, since these choices allow for many desirable properties to be satisfied by the resulting $\Delta^{d_T,\Sigma}$ and $\Delta^{d_H,\Sigma}$, see Table 4.15.

The unanimity-principle is not satisfied by $\Delta^{d_T,\Sigma}$ and $\Delta^{d_H,\Sigma}$, however this is not a bad thing in this context. If the unanimity on an issue is not respected by the collective decision, this is due to the rest of the judgments. It is more important that the initiator selects the "right" decision than preserve unanimity. For the same reason, majority-preservation is not a required property either. However, one nice feature of this property is that it allows the initiator to fast determine, in linear time with respect to number of agents and number of cues, the collective decision by checking if the majority is consistent. In addition, the d_H/d_T and $\sum$ selection has good computational-theoretic properties, particularly when no weights are given to the judgments, which we showed in Section 3.6.

Although the agents may abstain on any of the agenda issues, the collective set of judgments must contains as little abstentions as possible. In particular, the collective judgment regarding the goal should not be $\frac{1}{2}$, because this judgment leaves the initiator with no decision whether the goal is an adequate solution for the problem or not. Consequently, the judgment aggregation rule used must be such that all the sequences that it selects are from the restricted domain $X \subset \mathbb{A}$ in which the judgment assigned to the goal is either 0 or 1. To achieve this, the initiator can use the rules $\Lambda^{d_T,\Sigma}$ and $\Lambda^{d_H,\Sigma}$ constraining the co-domain to X.

If the initiator does not elicit judgments on the goal, he can also use one of the premise-based procedures of Definition 46. The biased procedure is adequate for the brave initiator mode or when time constraints to reach a decision are particularly severe, while the unbiased procedure is for the cautious initiator. The premise-based procedures cannot be applied to every agenda, and for some profiles they will generate a value $\frac{1}{2}$ regarding the goal decision. However, the premise-based procedures do have the low complexity to their advantage.

5.4 The overpass rescue scenario revisited

In this section we revisit the overpass rescue example and show how our team can reason when faced with the same problems following the recognition-primed group decision model. We use Łukasiewicz logic for the judgment aggregation problems.

5.4.1 The sub-goal of securing the victim

The initiator is called to rescue a woman who either fell or jumped off a highway overpass, and instead of falling to her death, had landed on the metal supports of a highway sign and is dangling there when five executor agents arrive on the scene. The initiator is in remote communication with the executors. There are two executors, E_1 and E_2 on the overpass, and three in a hook-and-ladder truck, E_3, E_4, and E_5. From the description the initiator got when called to the rescue, he determines that the team has two concurrent goals:

g_1: save the woman and

g_2: prevent the woman's body from falling on a moving car on the highway below her.

Example 5.4.1. *As a first sub-goal the lieutenant considers:*

s_1*: team members climbing up to the woman to secure her.*

The initiator will have the team adopt sub-goal s_1 if the cue, (c_1), at least two agents can get a good brace against the struts *is present, and if the agents think that $c_1 \rightarrow s_1$ is the case. The cue $c_1 \rightarrow s_1$ is the opinion that the s_1 can be accomplished if c_1 is present. The initiator assesses the situation, by requesting information from the executors,* i.e.*, their opinions on whether $\{c_1, c_1 \rightarrow s_1\}$ are the case or not. The replies of the executors are given in Table 5.2.*

The initiator uses the profile π and the weight matrix W in which each each agent is assigned the weight $w = 9.3 - dfv$, where dfv is the distance from the victim. We do not assign weights for the judgments on s_1 since these are deduced. The rationale is, the closer the agent is to the victim, the more reliable his judgment is.

Agents	c_1	$c_1 \to s_1$	s_1	Distance from victim
$\{E_1, E_2\}$	yes	yes	yes	1m, 1.3m
$\{E_3, E_4, E_5\}$	no	?	?	7m, 8m, 8.3m

Table 5.2: Contributed opinions and deduced goal opinions regarding cues on s_1.

$$\pi = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \qquad W = \begin{bmatrix} 8.3 & 8.3 & 1 \\ 8 & 8 & 1 \\ 2.3 & 2.3 & 1 \\ 1.3 & 1.3 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

The following collective decisions are obtained, with respect to the different rules used. X is the set of all consistent sequences in which the decision on s_1 is either 0 or 1.

$$\begin{aligned} \Lambda^{d_T,\Sigma}(\pi, W, X) &= \langle 1,1,1 \rangle \\ b - pbp(\pi, W) &= \langle 1,1,1 \rangle \\ u - pbp(\pi, W) &= \langle 1,1,1 \rangle \end{aligned}$$

Therefore the goal s_1 is adopted.

The tentative plan for s_1 the executer proposes is that E_1 grabs the legs of the victim and E_2 her shoulders, while the rest of the agents execute a traffic blocking procedure. He informs the agents about this plan. All of the agents approve it and the initiator informs all that this is the plan to be executed. The agents execute their corresponding actions and inform the initiator when they are done.

5.4.2 The sub-goal of extracting the victim

As a another sub-goal, the initiator considers how to pull the woman to safety. He first considers the solution

s_2: use a rescue harness.

Based on his experience, the initiator deems s_2 a satisfactory solution if and only if at least one of the following cues are the case:

c_2: the victim is in a sitting position,

c_3: the victim is in a face-up position,

c_4: the victim can safely be shifted in a sitting up position or in a face-up position,

c_5: the harness can be attached from the back without hurting the victim.

The initiator requests the executors opinions regarding these cues, and he further specifies that the cues are subject to the constraint $(c_2 \vee c_3) \to c_4$. The constraints encodes the “obvious” information that if a victim is already sitting or facing up, then c_4 is trivially the case. The opinions of the executors are given in Table 5.2.1.

The initiator uses the profile π and the weight matrix U .

$$\pi = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

The following collective decisions are obtained. X is the set of all consistent sequences in which the decision on s_2 is either 0 or 1. For the premise-based procedures d_T is used as well.

$$\begin{aligned} \Lambda^{d_T,\Sigma}(\pi,W,X) &= \langle 0,0,0,0,0\rangle, \langle 0,0,0,1,1\rangle \\ b-pbp(\pi,W) &= \langle 0,0,0,0,0\rangle \\ u-pbp(\pi,W) &= \langle 0,0,0,\tfrac{1}{2},\tfrac{1}{2}\rangle \end{aligned}$$

If the initiator used the rule $\Lambda^{d_T,\Sigma}(\pi,W,X)$ he needs to break the tie between the decisions $\langle 0,0,0,0,0\rangle$ and $\langle 0,0,0,1,1\rangle$. Since the tie is essentially between adopting the goal or not, he can do so by looking at how many of the deduced judgments for the goal are for and how many are against this goal. In this case he will proceed with $\langle 0,0,0,0,0\rangle$. Using the rule $u-pbp(\pi,W)$ it can be deduced that the initiator cannot decide if s_2 is a good solution. Let us assume that the initiator does not want to risk the victim and concludes that s_2 is not a satisfying solution. He generates another possible solution and now considers

s_3: use a rescue strap.

The necessary and sufficient conclusions for s_3 to be adopted are $c_2 \vee c_5$. He already has the group decision regarding these two cues and does not need to ask for them again. He uses the constraint $c_2 \vee c_5$ and determines that s_3 is an unsatisfactory solution and proceeds to generate another one. He now considers to

s_4: use a ladder belt.

This new sub-goal is a good solution when all of the following cues are present:

c_6: an agent can climb up the ladder (of the hook-and-ladder truck),

c_7: the ladder belt can be sledded under the woman and buckled from behind,

c_8: a rope can be tied to the snap,

c_9: the woman can be lifted by two agents.

The initiator elicits the opinions of the agents regarding these cues. For the cues c_6, c_7 and c_8 he only needs the opinion of the agents E_3, E_4 and E_5, since these are the ones that can operate the hook-and-ladder truck equipment. The opinions he gets are presented in Table 5.4.2.

The initiator uses the profile π and the weight matrix W .

Agents	c_6	c_7	c_8	c_9	s_4
$\{E_1\}$	–	–	–	no	–
$\{E_2\}$	–	–	–	yes	–
$\{E_3, E_5\}$	yes	yes	yes	yes	yes
$\{E_4\}$	yes	yes	yes	?	?

Table 5.3: Contributed opinions and deduced goal opinions regarding cues on s_4.

$$\pi = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 1 & \frac{1}{2} \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \qquad W = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

The following collective decisions are obtained. X is the set of all consistent sequences in which the decision on s_4 is either 0 or 1.

$$\Lambda^{d_T,\Sigma}(\pi, W, X) = \langle 1, 1, 11, 1 \rangle$$
$$b - pbp(\pi, W) = \langle 1, 1, 1, 1, 1 \rangle$$
$$u - pbp(\pi, W) = \langle 1, 1, 1, 1, 1 \rangle$$

The initiator announces that s_4 is adopted. He proceeds to formulate the tentative plan for s_4. His plan is to order E_3 to fetch the ladder belt and the rope, and both him and E_5 to position himself on the ladder platform. Afterwards E_4 is to lift the ladder. Once the belt is attached to the woman by E_3 or E_5, the agent's E_1 and E_2 lift the woman. (This way, if the woman falls he will be attached to the ladder). The initiator proposes this plan to the agents.

Agent E_4 object the plan stating that he is not able to operate the ladder. The initiator takes this constraint into account and proposes another tentative plan, same as the previous, except now the roles of E_4 and E_5 are switched. The new plan is announced to the agents and they confirm their agreement to it.

The initiator issues the orders with respect to the approved plan, and sets the moment t_1, after the belt is clasped and the rope tied and the woman had been lifted for couple of centimeters, as the time to re-evaluate the goal, cues and plan.

5.4.3 Verifying expectancies

At t_1, the agents E_1 and E_2 have relinquished holding the woman and attempted to lift her by pulling the rope. The agents need to report any perceived difference with respect to the estimate for the sub-goal they are currently pursuing, *i.e.,* s_4. All agents signal that all cues are present, but the sub-goal is not being accomplished since the belt slips from the woman. The initiator revises his cue-goal pattern and includes, as a necessary condition for adopting s_4, the cue

c_{10}: the belt is tightly clasped around the victim.

The agents are unanimous that c_{10} is not the case and the initiator announces that s_4 is no longer pursued.

He now comes up with another solution

s_5: use a rope tied around the waste of the victim.

The goal s_5 is a good solution if, apart from cues c_6 and c_9, also the cue c_{11}

c_{11}: a rope can be tied to the victim,

is the case. The initiator asks agents E_3, E_4 and E_5 for their opinions on whether c_{11} is the case. They unanimously judge c_{11} as true and the initiators announces that s_5 is adopted (all of the rules $\Lambda^{d,T}, b-pbp$ and $u-pbp$ satisfy unanimity). He now devises a plan for s_5, which consists of the agents on the ladder platform tying one end of the rope to the waist of the victim and throw the other end to the agents E_1 and E_2. The agents accept this plan and proceed with its execution.

5.5 Revision of emerging states

"Members of adaptive teams utilize their pooled resources (*i.e.*, knowledge gained from learning) to adjust their actions according to situation requirements" (Burke et al., 2006, pg. 1190). The adaptation in our model is executed through verifying expectancies and through revising the situation assessment agreements when new information becomes available.

The collective values assigned to the cues are an estimate of what the state of the world is, hence the value can be confirmed or refuted by later observations. For instance, as the executors proceed with executing the plan for goal s_1, they confirm the estimate on c_1, namely they can get a good brace against the struts. If instead the agents find that c_1 is impossible, the agreement on c_1 would be in contradiction with the observation. Information contradicting the agreed value of a cue may lead to the change in the decision to pursue a goal.

After a plan is adopted and execution starts, an agent may report a constraint regarding the plan. This initiator uses the constraint to adapt the plan. All the executors are informed of the change. A cause for a plan revision is also the revision in the situation assessment, *i.e.,* after the adopted judgment sequence A had been revised.

Regardless of whether the update is scheduled, *e.g.,* , after a task is executed, or caused by a new observation, there are two types of information that can be cause for revision: *a new constraint* on an agreed plan, or a *a cue value being determined*. The revision of agreements depends on whether the opinions regarding the goal(s) were explicitly elicited or reached as a deduction from the cue judgments. We first consider the case when the opinions on the goals were deduced.

Assume that the agenda is $\mathcal{A} = \mathcal{A}^c \cup \mathcal{A}^g$, where $\mathcal{A}^c$ is the set of cues and $\mathcal{A}^g$ a set of, corresponding, accepted goal(s). Let A be the sequence of agreed values for $\mathcal{A}$, according to which $A(g) = 1$, and let $v(a)$ be the observed value of a. The initiator needs to revise the sequence A so that it contains the observed value $v(a)$ while remaining consistent.

The value $v(a)$ can be established in two ways. If a is subject to a scheduled verification, then the executors have an opportunity to re-state their judgments on a. The initiator elicits these judgments and obtains $v(a)$ by comparing $n_w^+(a)$ with $n_w^-(a)$. Recall that $n_w^+(a)$ denotes the sum of weights of agents who judge an issue $a \in \mathcal{A}$ to be true, with respect to some profile

π and weight matrix W; $n_w^-(a)$ denotes the sum of the weights of the agents who judge a to be false. Observe that the weight of an agent regarding an issue can change between two elicitations. The second way to establish $v(a)$ is when an executor directly observes it. A revision of A follows.

The revision is a rule $\circledast$ that assigns a consistent judgment set to a judgment set A given the new value of an agenda issue, *i.e.*, $\circledast : \mathbb{A}(\mathcal{A}, \mathcal{R}, \models) \times \mathcal{A} \times T \mapsto \mathcal{P}(\mathbb{A}(\mathcal{A}, \mathcal{R}, \models))$. We can state the desirable properties for $\circledast$, motivated by the need to minimize the resources spent on this process.

The first property is that a revision should be done only when necessary. Let $T = \{0, \frac{1}{2}, 1\}$ be the set of values with $\frac{1}{2}$ denoting the abstention. The revision does not need to be done if $A(a) = v(a)$.

- $\circledast(A, v(a)) = A$ when $A(a) = v(a)$. (Estimates verified)

The revision should be *prioritized*, namely after the revision, $A^*(a) = v(a)$, where A^* is the revised judgment sequence. The observed value can only be $v(a) = 0$ or $v(a) = 1$, but $v(a) = \frac{1}{2}$ can be obtained by pooling the agent's opinions. In the case of $v(c) = \frac{1}{2}$, revision is not needed since $v(c) = \frac{1}{2}$ does not increase the knowledge of the initiator and a sequence cannot be made inconsistent by replacing a judgment in it with $\frac{1}{2}$. On the contrary, the more abstentions there are in the agreement, the more difficult it is for the initiator to establish the course of action.

- $\circledast(A, v(a) = \frac{1}{2}) = A$. (No increase of information)
- If $\circledast(A, v(a)) = A'$, then $A'(a) = v(a)$ for $v(a) \neq \frac{1}{2}$. (Success)

A potentially desirable property of revision can be *stability*. If the new information can be consistently embedded in the old agreed judgment sequence, then the reset of the judgment sequence should not be changed. This is the property of stability. Let $\hat{a}$ denote the set corresponding to $v(a)$: $\hat{a} = \{a\}$ if $v(a) = 1$, $\hat{a} = \{\neg a\}$ if $v(a) = 0$ and $\hat{a} = \emptyset$ if $v(a) = \frac{1}{2}$.

- If $(\hat{A} \backslash \{a, \neg a\}) \cup \hat{a} \cup \mathcal{R} \not\models \bot$ then $\circledast(A, v(a)) = (\hat{A} \backslash \{a, \neg a\}) \cup \hat{a}$. (Stability)

However, this property can lead to the initiator revising into a sequence that is not constructive, as it can be illustrated through an example.

Example 5.5.1 (Revising an agreement)**.** *Consider the goal s_2 and assume that this is the one the agents agreed on pursuing. The agenda is $\mathcal{A} = \{c_2, c_3 c_4, c_5, s_2\}$ and $\mathcal{R} = \{(c_2 \vee c_3 \vee c_4 \vee c_5) \rightarrow s_2, (c_2 \vee c_3) \rightarrow c_5\}$; with agreement reached $A = \langle 1, 0, 0, 1, 1 \rangle$. Let $v(c_2) = 0$ be an observed new value. If the initiator revises with a rule that satisfies stability, then the obtained revision is $A' = \langle 0, 0, 0, 1, 1 \rangle$. However, recall that $(c_2 \vee c_3) \rightarrow c_5$ hence it cannot be known whether the agents, without $c_2 = 1$ would have the opinion $c_5 = 1$. This information will surface once the initiator alters the plan and the agents try to execute the specified actions.*

Stability is still desirable if the cue a on which new information is observed is logically independent from the other cues. This property is *stability of independent cues*. Let A^c denote the sub-sequence of A containing the judgments on the cues. A cue a is *independent* from the rest of the cues in $\mathcal{A}^c$ if and only if $\hat{A}^c \backslash \{a, \neg a\} \not\models \hat{a}$ is true.

Assume that the initiator in Example 5.5.2, asked the opinions of the agents regarding c_5 only. The question is how are these new opinions to be integrated in A. One way is to compare $n_w^+(c_5)$ with $n_w^-(c_5)$ and embed the resulting value in A'. Another is to ask the agents to apply the rule $\circledast$ individually and report their revised judgment sequences. When the opinions of the agents regarding the goal are not deduced, but elicited explicitly, the agents should be given a chance to revise them individually as well. The initiator aggregates the new sequences, using the same aggregation rule as in the first aggregation. As a consequence of the stability of independent cues property, the agents whose initial judgments were confirmed by the new information do not revise their sequence.

Example 5.5.2 (Revising by re-aggregation)**.** *Assume that the original profile of judgments for goal s_2 was π. Applying the revision individually, agents $E_3 - E_5$ do not change their judgments since they had already judged c_2 as false. A possible new profile is thus π^*.*

$$\pi = \begin{matrix} E_1,E_2 \\ E_3,E_4 \\ E_5 \end{matrix} \overset{\begin{matrix} c_2 & c_3 & c_4 & c_5 & s_2 \end{matrix}}{\begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & \frac{1}{2} & 1 & \frac{1}{2} & 1 \end{bmatrix}} \qquad \pi^* = \begin{matrix} E_1,E_2 \\ E_3,E_4 \\ E_5 \end{matrix} \overset{\begin{matrix} c_2 & c_3 & c_4 & c_5 & s_2 \end{matrix}}{\begin{bmatrix} 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & \frac{1}{2} & 1 & \frac{1}{2} & 1 \end{bmatrix}}$$

Figure 5.4: The profile π and a possible revision π^*.

The revision is a rule that can select several judgment sequences as an outcome. In this case, the A' that contain a value of the goal $A(g) = 1$ are preferred, since in this case the initiator does not need to generate a new goal.

5.6 Conclusions

In this chapter we study decision-making for a hierarchical team in an uncertain environment. We construct a model by lifting the recognition-primed decision (RPD) model from an individual agent model to a group model. According to the RPD model, satisficing decisions are identified by identifying a set of relevant cues and verifying whether these cues are present or absent. The RPD model and our extension are applicable only when the agent responsible for the decision is familiar with the problem to which the decision is related.

In our team, the agent responsible for making the group decision uses a judgment aggregation rule to aggregate the opinions of the members without considering his own. We considered as an illustration a case scenario collected by (Klein, 1999).

Given the hierarchical nature of the team, the ternary value-type of the aggregated judgments and the presence of weights we propose that the rules of Chapter 3 are used for aggregation within the scope of the recognition-primed group decision model we propose, in particular the $\Delta^{d_H,\Sigma}$ and $\Delta^{d_T,\Sigma}$. In the cases when the initiator only considers the judgments with respect to cues, but not regarding the goal, he can use the extended premise-based rules. Due to the time constraints posed on the initiator, the rules $\Lambda^{d_T,\Sigma}$ and $\Lambda^{d_H,\Sigma}$ are a good choice as well, since they allow the initiator to constrain the domain to permit only binary judgment sequences as the result of the aggregation.

In Section 3.6 we studied the winner determination problem for the rules $\Delta^{d_H,\Sigma}$ and $\Delta^{d_T,\Sigma}$. The winner determination problem is a decision problem, it only tells us how efficiently can we confirm, for the worst-case scenario of a profile and weight matrix, that a particular judgment sequence is a result of the aggregation for a particular aggregation rule. However, the important efficiency analysis in the case of resource constrained agents that use the recognition-primed group decision model is the search version of the winner determination problem: how efficiently can a sequence that is the result of the aggregation for a particular rule *be found*? To answer this question we need to analyze the functional complexity of the winner determination problem, particularly for the rules $\Delta^{d_H,\Sigma}$ and $\Delta^{d_T,\Sigma}$.

Chapter 4 we studied the properties of the $\Delta^{d_H,\Sigma}$ and $\Delta^{d_T,\Sigma}$, however we did include the extended premise-based rules and the domain-restricted $\Lambda^{d,\odot}$ rules in our analysis. Which of the social-theoretic properties defined and studied in Chapter 4 continue to apply when the co-domain is restricted? For which co-domain restrictions are properties restored or fail? We encountered a counter example, the decision on goal s_2 in Section 5.4.2, that confirms that although $\Delta^{d_T,\Sigma}$ is majority-preserving, its counterpart $\Lambda^{d_T,\Sigma}$ is not.

We considered a mixed human-robot team in which there is one human, the initiator, which has the role of a leader. The initiator is not on the ground where the problem is and coordinates remotely with the ground agents, *i.e.*, the executors. The model we developed heavily depends on the experience and creativity of the initiator, which is why this agent is human. It is the initiator who matches the problem with the corresponding goal and the goal with corresponding cues. He elicits and aggregates the opinions, generates the plan and implements the revision. An executor only needs to be able to form a judgment regarding a cue when asked for one, and evaluate whether a given action sequence is within his capacities.

How difficult would it be for a group of purely artificial agents to reason according to the team recognition-primed decision model? We can abstract an experience of a human initiator to a *case*. A case is the product of learning. The case that the lieutenant constructs after he considers the events of the overpass rescue example can be modeled as on Figure 5.5. Each case can be modeled as a quadruple $case = \langle p, G, C, \mathcal{R} \rangle$, consisting of the encountered problem p, the set of goals G pursued to solve it, the set of cues C which identified the goal(s) as adequate and the relational structure $\mathcal{R}$ for the goal and cues.

A human commander gathers cases from personal experience, but also by exchanging experiences with colleagues. A set of cases can be supplied to an (artificial) initiator agent. However, the power of the human commander is in the ability to recognize cases as similar. A successful non-human initiator must be able to do the same: perform a swift search through the cases and identify the case most similar to the current situation. The similarities can be found between problem characteristics, but also in cues.

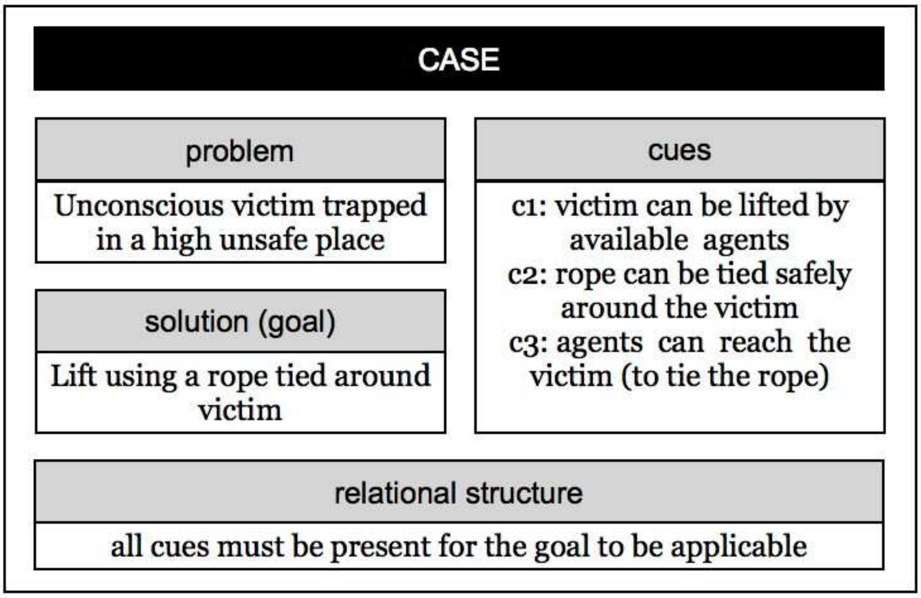


Figure 5.5: The case constructed from the overpass rescue example.

6

Group intentions are social choice with commitment

Abstract. In this chapter we consider the problem of forming group intentions as an example of a consensual group decision-making problem. An agent intends *g* if it has chosen to pursue goal *g* and is committed to pursuing *g* . How do groups decide on a common goal? Social epistemology offers two views on collective attitudes: according to the summative approach, a group has attitude *p* if all of the group members have the attitude *p*; according to the non-summative approach, for a group to have attitude *p* it is required that the members together agree that they have attitude *p*. The summative approach is used extensively in multi-agent systems. The main advantage of this approach is the simplicity of determining if all group members have the same attitude. The main disadvantage is that it does not allow for groups that can reach agreement to act together, which is why it has been heavily criticized in the social epistemology literature. We propose a formalization of non-summative group intentions, using judgment to determine the group goals. We use judgment aggregation as a decision-making mechanism and a multi-modal multi-agent logic to represent the collective attitudes, as well as the commitment and revision strategies for the groups intentions.

6.1 Introduction

An intelligent agent interacts with its environment and other agents. This interaction includes cooperation. In order for the agents to cooperate they need to establish what are their group goals, and subsequently intentions. Of all the collective attitudes, the formation of group intentions is possibly both most interesting and challenging since an intention is inevitably related to other attitudes. Cohen and Levesque (1990) argue that the intention of an agent is the goal that he chooses to pursue and is committed to pursuing, and it has been argued since that the goals of an agent are intricately linked with the agent's beliefs (Castelfranchi and Paglieri, 2007; Boella et al., 2007).

How collective attitudes are formed and what is their nature is studied by *social epistemology*. There are two predominant views, the summative and the non-summative view, regarding the relation between the attitude of the group and the corresponding attitudes individually held by the members. According to the *summative view*, a group has attitude *p* if all or most of the group members have the attitude *p* (Quinton, 1975; Hakli, 2006). According to the *non-*

summative view a group has an attitude *p* if the members *together agree* that they have that attitude *p* (Meijers, 2002; Gilbert, 2002, 2009).

Within the context of multi-agent systems, the concept of collective intentions is studied and formalized in (Dunin-Keplicz and Verbrugge, 2010, Chapter 3) and also in (Singh, 1990; Jennings, 1995; Grosz and Hunsberger, 2007). In (Dunin-Keplicz and Verbrugge, 2010, Section 3.9), we find a detailed overview of the various formalizations of group intentions. It can be observed that all these formalizations follow the summative view on collective intentionality. As observed in (Dunin-Keplicz and Verbrugge, 2010, Section 3.9), collective intentions and collective commitments appear as central in the work of Margaret Gilbert (Gilbert, 1987, 2002, 2007, 2009), who upholds the non-summative view of intentions, but whose work is predominately philosophical.

The advantage of the summative approach is that it is very easy to determine when a group goal exists, particularly in hierarchical groups. In this case the agent responsible for producing the group decision only needs to confirm that no-one is of a different opinion. However, if a group acts only when everyone in the group is in unanimous agreement, then the situations in which the group can act are limited. For one, it is not likely that a very large number of agents, or a group of heterogeneous agents, would be always in unanimous agreement. This forces the size of feasible groups to be kept small and/or the group to be kept homogenous.

Under the summative approach, when the goal of the group is established there is no ambiguity regarding what the individual goals of the agents are, implying that the agents are perfectly cooperative regarding each goal they undertake. Not all groups are perfectly cooperative since often agents undertake group goals while pursuing individual goals of their own. According to the non-summative approach, from the existence of a group goal *g* it cannot be deduced what the individual goals of the agents are. This allows not only purely cooperative groups to be modeled.

We can use the concepts devised by the work of Gilbert to formalize collective attentions in a new way. In this chapter we formalize non-summative group intentions and joint commitments.

How can a group agree on what its intentions are?

Following the paradigm of "intention is choice with commitment" we need to discover how a group can decide which goals to pursue and also how can it commit to pursuing them?

A rational agent makes decisions based on what he believes, what he knows and what he desires. Each group member can express whether he is for or against a candidate group goal, but also how his opinions and knowledge support and justify his goal choice. We need a mechanism for generating group goals that aggregates individual opinions into collective attitudes. A group that jointly decided on a course of action is jointly committed to uphold that decision (Gilbert, 2007).

In practical reasoning, the roles of intentions can be summarized as: intentions drive means-end-reasoning, intentions constrain future deliberation, intentions persist long enough, according to a reconsideration strategy, and intentions influence beliefs upon which future practical reasoning is based (Dunin-Keplicz and Verbrugge, 2010). A formalization of group intentions should be completed with a formalization of group intention persistence and reconsideration strategies. These strategies are difficult to develop when the decision to pursue (or not) a goal is devoid from the knowledge and beliefs that rationalize and justify it. Therefore we need a group agreement on not only whether to pursue a goal but also why to pursue or reject it. In consequence we need the agents to express not only if they want the goal to be pursued, but also the reasons, stemming from their individual beliefs and knowledge, that justify their view on the goal.

Our research question thus breaks down into the following sub-questions:
1. How to aggregate the individual opinions into group beliefs and goals?
2. How to represent individual opinions and non-summative group attitudes?
3. How can groups persist in their intentions?
4. How can groups reconsider their attitudes?

The relation between individual goals and beliefs can be specified and analyzed in modal agent logics like BDI_{LTL} (Schild, 2000). The challenge is to find an adequate representation for the individual opinions and the non-summative beliefs, goals and intentions into multi-agent logic. We give an extension logic AGE_{LTL} that fuses existing modal logics to provide the adequate modalities. We use this logic to represent the group intention and reconsideration strategies.

We require that the group has a set of candidate group goals, a relevance order over this set, as well as a set of constraints, one for each candidate goal, in the form of logic formulas, that express what is the relation between a goal and a given set of reasons. The members are required to have the ability to form and communicate "yes" or "no" judgments regarding a candidate goal and associated reasons.

We need a mechanism for generating group beliefs and goals that aggregates individual opinions into collective attitudes. Since the agents express their opinions regarding a set of logically related issues, beliefs and opinions on whether a goal is to be pursued, a judgment aggregation rule, such as the ones we considered in Chapters 2 and 3 is an adequate bases for such a mechanism.

A non-summative goal needs the agreement of all agents to be established. Consequently, there is no one agent that can be responsible for the group decision, and it can be consistently assumed that the agents' opinions are all considered as of equal weight in the aggregation. An agreement must be responsive to the opinions of the group members and satisfy such properties as majority-preservation and unanimity which we presented in Chapter 4.

Cohen and Levesque (1990), proclaim that intentions are choice (of a goal) with commitment. Judgment aggregation is a social choice mechanism. Following the intuition of Cohen and Levesque, (a non-summative) group intention is (a group goal determined by) social choice with commitment.

The layout of the chapter is as follows. In Section 2 we discuss how to choose group goals. We first summarize the non-summative view on collective attitudes. We then extend BDI_{LTL} with the necessary modalities for representing these group attitudes and the concepts from judgment aggregation. We introduce a judgment aggregation framework using this logic extension and, in Section 6.3, show how it can be used. Sections 6.4 and 6.5 respectively study the commitment and reconsideration strategies. Related work, conclusions and outlines for future work are in Section 6.6.

6.2 Non-summative group attitudes formalized

First we discuss how non-summative goals and beliefs are determined and then introduce the logic AGE_{LTL} which is used for representing these attitudes. The formal model of judgment aggregation, using this logic, is given in Section 6.2.3.

6.2.1 From individual opinions to group attitudes

The intention of the group is formalized using the summative approach, according to existing theories such as (Levesque et al., 1990; Jennings, 1995; Dunin-Keplicz and Verbrugge, 2010), following (Bratman, 1993) and (Rao et al., 1992): "g is the intention of the group" is equivalent to "g is the individual intention of all the group members". Unlike the joint intention of, for example (Dunin-Keplicz and Verbrugge, 2010), our group intention is not necessarily decomposable into individual intentions: "an adequate account of shared intention is such that it is not necessarily the case that for every shared intention, on that account, there be correlative personal intentions of the individual parties" (Gilbert, 2009, pg.172).

Example 6.2.1. *Let $C = \{w_1, w_2, w_3\}$ be a crew of cleaning robots. We denote the group goal to clean the meeting room with g_1, and the reasons to adopt this goal with: there are no people in the room (p_1), the room is dirty (p_2), the garbage bin in it is full (p_3).*
The individual beliefs of the robots on whether g_1 should be the group goal are justified by individual beliefs on p_1, p_2, p_3 using the constraint $(p_1 \wedge (p_2 \vee p_3)) \leftrightarrow g_1$.

The group goal Gg_1 is not necessarily decomposable into individual goals g_1 upheld individually by the agents. Assume that robot w_1 in Example 6.2.1 is a mopper, the robot w_2 is a garbage collector and the robot w_3 sprays adequate cleaning chemicals. It can be that the individual goals of w_1 and w_2 are to clean the room. The goal of w_3 may be others, but the group agreed to pursue g_1 and it, being committed to g_1 as part of the group, will spray the cleaner as an act towards accomplishing g_1.

We formalize only goals that can be achieved by the group as a whole. Whether these goals can be achieved by joint actions or by a combination of individual actions is out of the scope. We define group intention to be the goal, which the members agreed on, and by that, are committed to pursuing.

The robots in Example 6.2.1 can disagree on various issues when reaching a decision for a group goal. Assume that one robot believes the room is occupied and thus, according to it, the group should not pursue g_1. According to the other two robots, the group should pursue g_1. The second robot is of the opinion that the garbage bin is full and the floor is clean, while the third believes that the floor is dirty. According to the non-summative view of collective beliefs, a group believes p if the group members together agree that as a group they believe p. The question is: how can a judgment aggregation rule be used to aggregate the beliefs of the robots?

To use judgment aggregation for aggregating the opinions of the robots, one needs to represent the individual and collective judgments as logic formulas. A logic of belief-desire-intention, a modal logic with modal operators B_i for belief of agent i, D_i for desire and I_i for intention, is insufficient to model these doxastic attitudes. According to Gilbert, "it is not logically sufficient for a group belief that p either that most group members believe that p, or that there be common knowledge within the group that most members believe that p"

(Gilbert, 1987, pg.189). Furthermore, "it is not necessary that any members of the group personally believe p" (Gilbert, 1987, pg.191). A w_1 robots judgment "yes" on $\neg p_1$ is not implied by nor it implies that robot's belief $B_{w_1}\neg p_1$.

Hakli (2006) summarizes the difference between beliefs and acceptances as: (1) beliefs are involuntary and acceptances are voluntary; (2) beliefs aim at truth and acceptances depend on goals; (3) beliefs are shaped by evidence and acceptances need not be; (4) beliefs are independent of context and acceptances are context-dependent; and (5) beliefs come in degrees and acceptances are categorical. We find that an individual judgment is closer to an acceptance than to a belief because like acceptances, judgments are voluntary, they depend on goals and are context-depend. Like beliefs, judgments are also are shaped by evidence. For these reasons we choose to represent judgments as acceptances.

There is a debate among social epistemologists on whether collective beliefs are proper beliefs or they are in essence acceptances (Gilbert, 2002; Meijers, 2002; Hakli, 2006). Since we use acceptances for individual judgments, we deem most adequate to use acceptances to represent the collective judgments as well.

The set of collective acceptances is the agreed upon group goal and group beliefs. Having group beliefs in support of group goals is in line with (Castelfranchi and Paglieri, 2007) who argue that the goals should be considered together with their supporting "*belief structure*". In Example 6.2.1, the constraint $(p_1 \wedge (p_2 \vee p_3)) \leftrightarrow g_1$ is nothing else but the "*belief structure*" for g_1. We use the group beliefs to define commitment strategies in Section 6.4.

6.2.2 The logic AGE_{LTL}

The logic we introduce to represent non-summative group attitudes is a fusion of two K-modal logics (Chellas, 1980), the logic of acceptance (Lorini et al., 2009) and the linear temporal logic (Pnueli, 1977). As such, it inherits the decidability properties of the fused logics (Wolter, 1998). The syntax of AGE_{LTL} is presented in Definition 82. The semantics is as that given by (Schild, 2000) for BDI_{CTL}.

To model the considered group goals we use a single K modal operator G. Thus Gg, where g is a propositional formula, is to be interpreted as "*g is a group goal*". Since we are interested in modeling the change upon new information, we also need to model these observations of new information. To this end we add the K modal operator E, reading $E\varphi$ as "it is observed that φ".

To model the individual and collective judgments we use the modal operator of acceptance A_S, where S is a subset of some set of agents N. $A_S\varphi$ allows us to represent both individual judgments, $S = \{i\}$, for $i \in N$ and collective judgments with $S = N$.

Definition 82 (Syntax)**.** *Let N be a non-empty set of agent names, with $S \subseteq N$, and L_P be a set of atomic propositions. The admissible formulae of $AGE_{LTL}$are formulae ψ_0, ψ_1 and ψ_2 of languages $\mathcal{L}_{prop}$, $\mathcal{L}_G$ and $\mathcal{L}_{AE_{LTL}}$ correspondingly, given here in BNF form:*
$\psi_0 ::= p \mid (\psi_0 \wedge \psi_0) \mid \neg\psi_0$
$\psi_1 ::= \psi_0 \mid G\psi_0$
$\psi_2 ::= \psi_0 \mid A_S\psi_1 \mid E\psi_2 \mid \boldsymbol{X}\psi_2 \mid (\psi_2\boldsymbol{U}\psi_2)$

*The p ranges over L_P and S over $\mathcal{P}(N)$. Moreover, $\Diamond\varphi \equiv \top\boldsymbol{U}\varphi$, $\Box\varphi \equiv \neg\Diamond\neg\varphi$, and $\varphi\boldsymbol{R}\varphi' \equiv \neg(\neg\varphi\boldsymbol{U}\neg\varphi')$. **X**, **U** and **R** are standard operators of LTL. We recall the reader of the semantic of the **X** and **U** later on in this section when we introduce the semantics of AGE_{LTL}.*

Example 6.2.2. *Consider Example 6.2.1. Gg_1 represents that cleaning the room is a group goal. $A_C Gg_1$ represents that the group C accepts cleaning the room as its group goal. $A_{\{w_3\}} g_1$ represents that agent w_3 is of the opinion that the group should adopt g_1. Ep_1 represents the observation that there are no people in the meeting room. $E\Box\neg p_1$ denotes that it is impossible to clean the meeting room.*

We use the linear temporal logic to model the change of group attitudes. By using LTL we do not need to distinguish between path formulas and state formulas. BDI_{LTL} uses, for example $B\Box a$ to quantify over traces. We can use E for that purpose.

We define the intention of the group of agents S to be their acceptance of a goal, where S ranges over 2^{Agt} as

$$I_S \psi \equiv_{def} A_S G \psi.$$

Semantics of AGE_{LTL}

As mentioned, the semantics of AGE_{LTL} follows the semantics of BDI_{LTL} presented in (Schild, 2000). A Kripke structure is defined as a tuple $\mathcal{M} = \langle W, \mathcal{R}, \mathcal{G}, \mathcal{E}, \mathcal{A}, L \rangle$. The set W is a set of possible situations. The set $\mathcal{R}$ is a set of pairs identifying the temporal relation over situations $\mathcal{R} \subseteq W \times W$. The set $\mathcal{G}$ is a set of pairs identifying the goal relation over situations $\mathcal{G} \subseteq W \times W$. Lastly, the set $\mathcal{E}$ is a set of pairs identifying the observation relation over situations $\mathcal{E} \subseteq W \times W$. The element $\mathcal{A}$ is a map $\mathcal{A} : 2^N \mapsto W \times W$. The mapping $\mathcal{A}$ assigns to every set of agents $S \in 2^N$ a relation $\mathcal{A}_S$ between possible situations. L is a truth assignment to the primitive propositions of L_P for each situation $w \in W$, i.e., $L(w) : Prop \mapsto \{true, false\}$.

Given a structure $\mathcal{M} = \langle W, \mathcal{R}, \mathcal{G}, \mathcal{E}, \mathcal{A}, L \rangle$ and $s \in W$, the truth conditions for the formulas of AGE_{LTL}(in a situation s) are:

- $\mathcal{M}, s \not\models \bot$;
- $\mathcal{M}, s \models p$ if and only if $p \in L(p)$;
- $\mathcal{M}, s \models \neg\varphi$ if and only if $\mathcal{M}, s \not\models \neg\varphi$;
- $\mathcal{M}, s \models \varphi \wedge \psi$ if and only if $\mathcal{M}, s \models \varphi$ and $\mathcal{M}, s \models \psi$;
- $\mathcal{M}, s \models A_S \varphi$ if and only if $\mathcal{M}, s' \models \varphi$ for all $(s, s') \in \mathcal{A}(S)$;
- $\mathcal{M}, s \models G\varphi$ if and only if $\mathcal{M}, s' \models \varphi$ for all $(s, s') \in \mathcal{G}$;
- $\mathcal{M}, s \models E\varphi$ if and only if $\mathcal{M}, s' \models \varphi$ for all $(s, s') \in \mathcal{E}$;
- $\mathcal{M}, s \models \mathbf{X}\varphi$ if and only if $\mathcal{M}, s' \models \varphi$ for the s', $(s, s') \in \mathcal{R}$
- $\mathcal{M}, s \models \varphi \mathbf{U} \psi$ if and only if $\mathcal{M}, s \models \varphi$; $\mathcal{M}, s^i \models \varphi$ for all s^i, $i \in \{1, 2, \ldots, k\}$ such that $\{(s, s^1), (s^1, s^2), \ldots (s^{k-1}, s^k)\} \in \mathcal{R}$ and for s^{k+1} such that $(s^k, s^{k+1}) \in \mathcal{R}$ it holds $\mathcal{M}, s^{k+1} \not\models \varphi$ and $\mathcal{M}, s^{k+1} \models \psi$.

A formula φ is true in a AGE_{LTL}model $\mathcal{M}$ if and only if $\mathcal{M}, s \models \varphi$ for every situation $s \in W$. The formula φ is valid (noted $\models_{AGE_{LTL}}$) if and only if φ is true in all AGE_{LTL}models. The formula φ is AGE_{LTL}-satisfiable if and only if the formula $\neg\varphi$ is not AGE_{LTL}valid.

For the purposes of constructing the formal judgment aggregation model, we emphasize that a set of sentences $M \subseteq AGE_{LTL}$ is called consistent if $M \not\models \bot$ and inconsistent otherwise. The logic AGE_{LTL} satisfies: for each pair $\{\varphi, \neg\varphi\} \in AGE_{LTL}$, $\{\varphi, \neg\varphi\} \models \bot$, and $\emptyset \not\models \bot$.

(C_1) For each set $\{a, \neg a\} \in AGE_{LTL}$ it holds $\{a, \neg a\} \models \bot$.

(C_2) Given a set $M \subseteq AGE_{LTL}$ such that $AGE_{LTL} \not\models \bot$, it holds that $M' \not\models \bot$ for every $M' \subset M$.

(C_3) For the empty set $\emptyset$ it holds that $\emptyset \not\models \bot$.

(C_4) For each set M such that $M \subseteq AGE_{LTL}$, there exists a superset $T \in AGE_{LTL}$ such that $T \not\models \bot$ and either $a \in T$ or $\neg a \in T$ for every pair $\{a, \neg a\} \in AGE_{LTL}$.

Axiomatization of AGE_{LTL}

In our logic we model only acceptances since the private mental states, such as beliefs, are modeled by the BDI_{LTL} logic which we extend. We include the axioms and the semantics for LTL, since we use LTL to define the commitment strategies of the agents in Section 6.4.

The modal operator $A_S\varphi$ we use is equivalent to the modal operator $A_{S:x}\varphi$ of the *acceptance logic* of (Lorini et al., 2009) with one syntactic and one semantic exception. These exceptions do not infringe on the decidability properties of the logic, as it can be observed by the decidability proof for acceptance logic provided in (Lorini et al., 2009).

The operator $A_{S:x}\varphi$ uses x ranging over a set of labels to describe the context under which the acceptance is made. In our case the context is that of the group and since we deal with only one group, we have no use of these labels. The context labels play no role in the semantics of the acceptance logic formulas.

On the semantic level, the axioms for $A_S\varphi$ are all the axioms of $A_{S:x}\varphi$ except two: the axiom inclusion ($Inc.$) and the axiom unanimity ($Un.$). Dropping ($Un.$) and ($Inc.$) does not affect the decidability of the logic of acceptance. ($Un.$)[1] states that if $A_{N:x}\varphi$, then $\forall i \in N$, $A_{\{i\}:x}\varphi$. In our case, it is the aggregation of individual acceptances that determines the collective acceptance and we do not require that the group accepting p entails that all the members accept p, a property of non-summative collective belief indicated by (Gilbert, 1987). The opposite property, *i.e.*, all the agents accepting p implies that the group accepts p, is ensured via the judgment aggregation mechanism. ($Inc.$) states that if a group C accepts φ, so will any subgroup $B \subset C$. In our case, the judgment aggregation over the input from group B can produce different group attitudes than the judgment aggregation over the input from a larger group C.

The axiomatization of the AGE_{LTL}logic is thus:

(**ProTau**) All principles of propositional calculus

(**LTLTau**) All axioms and derivation rules of LTL

(**K-G**) $G(\varphi \to \psi) \to (G\varphi \to G\psi)$

(**K-E**) $E(\varphi \to \psi) \to (E\varphi \to E\psi)$

[1] Not to be confused with unanimity introduced in judgment aggregation in Section 6.3

(**K-A**) $A_S(\varphi \to \psi) \to (A_S\varphi \to A_S\psi)$

(**PAccess**) $A_S\varphi \to A_M A_S\varphi$ if $M \subseteq S$

(**NAccess**) $\neg A_S\varphi \to A_M\neg A_S\varphi$ if $M \subseteq S$

(**Mon**) $\neg A_S\bot \to \neg A_M\bot$ if $M \subseteq S$

(**MP**) From $\vdash \varphi$ and $\vdash (\varphi \to \psi)$ infer $\vdash \psi$

(**Nec-A**) From $\vdash \varphi$ infer $\vdash A_S\varphi$

(**Nec-G**) From $\vdash \varphi$ infer $\vdash G\varphi$

(**Nec-E**) From $\vdash \varphi$ infer $\vdash E\varphi$

6.2.3 Agreeing on group intentions as a judgment aggregation problem

Our judgment aggregation model in AGE_{LTL} follows the judgment aggregation (JA) model in general logics of (Dietrich and List, 2007a) and Chapter 2.

We presume that all the goals which the group considers to adopt are given in a set of candidate group goals $\mathcal{G} = \{Gg \mid g \in \mathcal{L}_{prop}\}$. The decision problem of choosing or not a given group goal is specified by an agenda. The agendas here are pre-defined consistent sets of formulas representing an issue on which an agent casts his judgments. An agenda is *truth-functional* if it can be partitioned into premises and conclusions. In our case, the agendas consist of one conclusion, which is the group goal $g \in \mathcal{G}$ being considered. The relevant reasons for this group goal are premises.

Definition 83 (Agenda)**.** *An agenda $\mathcal{A} \subseteq \mathcal{L}_G$ is a consistent set of formulas, such that $\mathcal{A} = \mathcal{A}^p \cup \mathcal{A}^c$. The sets $\mathcal{A}^p$ and $\mathcal{A}^c$ are such that $\mathcal{A}^p \subseteq \mathcal{L}_{prop}$, $\mathcal{A}^c \subseteq \mathcal{L}_G$ and $\mathcal{A}^p \cap \mathcal{A}^c = \emptyset$.*

We remark that in judgment aggregation models, as the one of (Dietrich and List, 2007a), the distinction between conclusions and premises is only indicated by the partition but not formalized in the language of the agenda. The reason why we need a language more expressive than propositional logic to represent the agenda issues is that we want to explicitly formalize this distinction trough the modal operator G.

For a given agenda $\mathcal{A}$, each agent in the group N expresses his judgments by accepting (or not) the agenda issues. We define judgments formally in Definition 84.

Definition 84 (Judgment)**.** *Given a set of agents N and an agenda $\mathcal{A}$, for each issue $a \in \mathcal{A}$ the individual judgment of agent $i \in N$ is one element of the set $\{A_{\{i\}}a, A_{\{i\}}\neg a\}$. The collective judgment of N is one element of the set $\{A_N a, A_N\neg a\}$.*

The formula $A_{\{i\}}a$ is interpreted as agent i judges a to be true, while the formula $A_{\{i\}}\neg a$ is interpreted as agent i judges a to be false. Since the judgments are acceptances, we can assume that each agent is able to determine whether he accepts an issue or not. Consequently, a judgment $\neg A_{\{i\}}a$ is taken to be the same as judgment $A_{\{i\}}\neg a$, and the judgments $\neg A_N a$ the same as judgments $A_N\neg a$. In theory, an agent, or a group can also express the judgment of "do not know whether to accept a" via the formula $\neg A_{\{i\}}a \wedge \neg A_{\{i\}}\neg a$, or respectively $\neg A_N a \wedge \neg A_N\neg a$.

The goal and the reasons are logically related. These relations are represented by a set of constraints. In our model, we assume that the constraints are a set of formulas $\mathcal{R} \subseteq \mathcal{L}_G$. For each goal $Gg \in \mathcal{G}$ there is, provided together with the agenda, a set of constraints $\mathcal{R}_g \subseteq \mathcal{R}$. The set $\mathcal{R}$ contains all the constraints that the agent should observe when casting judgments. These constraints contain three types of information: rules describing how the goal depends on the reasons (justification rules $\mathcal{R}_g^{just}$), rules describing the constraints of the world inhabited by the agents (domain knowledge $\mathcal{R}_g^{DK}$) and rules that describe how g interacts with other candidate goals of the group (coordination rules $\mathcal{R}_g^{coord}$). Hence, the constrains for a group goal g are $\mathcal{R}_g = \mathcal{R}_g^{just} \cup \mathcal{R}_g^{DK} \cup \mathcal{R}_g^{coord}$.

We want the reasons for a goal to rationalize, not only the choice of a goal, but also its rejection. Having collective justifications for rejecting a goal enables the agents to re-consider adopting a previously rejected group goal. To this end, we require that the justification rules have the schema $Gg \leftrightarrow \Gamma$, where $\{Gg\} = \mathcal{A}_g^c$ and $\Gamma \in \mathcal{L}_{Prop}$ is a formula such that all the non-logical symbols of Γ occur in $\mathcal{A}_g^p$ as well.

The agents express their judgments on the agenda issues, but they accept the constraints *in toto*[2].

Example 6.2.3 (Example 1 revisited)**.** *Consider the cleaning crew from Example 6.2.1.* $\mathcal{R}_{g_1}^{just}$ *is* $(p_1 \wedge (p_2 \vee p_3)) \leftrightarrow Gg_1$ *and* $\mathcal{A}_{g_1} = \{p_1, p_2, p_3, Gg_1\}$. *Suppose that the crew has the following candidate group goals as well: place the furniture in its designated location* (g_2) *and collect recyclables from garbage bin* (g_3)*. The agendas are* $\mathcal{A}_{g_2} = \{p_4, p_5, p_6, p_7, Gg_2\}$, $\mathcal{A}_{g_3} = \{p_3, p_8, p_9, Gg_3\}$. *The justification rules are* $\mathcal{R}_{g_2}^{just} \equiv (p_4 \wedge p_5 \wedge (p_6 \vee p_7)) \leftrightarrow Gg_2$ *and* $\mathcal{R}_{g_3}^{just} \equiv (p_8 \wedge p_9 \wedge p_3) \leftrightarrow Gg_3$. *The formulas* $p_4 - p_9$ *are: the furniture is out of place* (p_4)*, the designated location for the furniture is empty* (p_5)*, the furniture has wheels* (p_6) *, the furniture has handles* (p_7)*, the agents can get revenue for recyclables* (p_8)*, there is a container for the recyclables* (p_9)*.*
An example of a domain knowledge could be $\mathcal{R}_{g_2}^{DK} \equiv \neg p_4 \rightarrow \neg p_5$*, since it cannot happen that the designated location for the furniture is empty while the furniture is not out of place. Group goal* Gg_3 *can be pursued at the same time as* Gg_1*, however,* Gg_2 *can only be pursued alone. Thus the coordination rule for all three goals is*
$\mathcal{R}_{g_1}^{coord} = \mathcal{R}_{g_2}^{coord} = \mathcal{R}_{g_3}^{coord} \equiv ((Gg_2 \wedge \neg(Gg_1 \vee Gg_3)) \vee \neg Gg_2)$.

To ensure that the judgments provided by the agents are usable for generating group goals, we impose that each of the individual judgments sets is complete and consistent as defined in Chapter 3.

Definition 85 (Admissible judgment set)**.** *Let* $\varphi = \{A_M\overline{a} \mid \overline{a} = a \text{ or } \overline{a} = \neg\, a, a \in \mathcal{A}\}$ *be the set of all judgments from agents* $M \subseteq N$ *for agenda* $\mathcal{A}$*. We define the set of accepted constraints* $\mathcal{R}_M = \{A_M r \mid r \in \mathcal{R}\}$*. The set of judgments* φ *is admissible if it satisfies the following conditions:*

- *for each* $a \in \mathcal{A}$*, either* $A_M a \in \varphi$ *or* $A_M \neg a \in \varphi$ *(completeness), and*
- $\varphi \cup \mathcal{R}_M \not\models \bot$ *(consistency).*

A profile, as in Chapter 2, is a set of judgment sets, *e.g.*,

[2]The agents accept the constraints as they are, as a whole set.

$\pi = \{A_{\{w_1\}}p_1, A_{\{w_1\}}p_2, A_{\{w_1\}}\neg p_3, A_{\{w_1\}}Gg_1, A_{\{w_2\}}\neg p_1, A_{\{w_2\}}p_2, A_{\{w_2\}}p_3, A_{\{w_2\}}Gg_1,$ $A_{\{w_3\}}p_1, A_{\{w_3\}}\neg p_2, A_{\{w_3\}}\neg p_3, A_{\{w_3\}}\neg Gg_1\}$ is a possible profile for Example 6.2.1.

We can also use the matrix notation

$$\pi = \begin{matrix} & p_1 & p_2 & p_3 & Gg_1 \\ w_1 & 1 & 1 & 0 & 1 \\ w_2 & 0 & 1 & 1 & 1 \\ w_3 & 1 & 0 & 0 & 0 \end{matrix}.$$

In the judgment sets in Chapter 2, one cannot distinguish whether a judgment set belongs to one agent or another. Using the acceptance operator to model judgments, we can make a distinction between the individual judgments. In judgment aggregation, the collective judgment set of a group of agents is obtained by applying a judgment aggregation function to the profile. The judgment aggregation rules F we use here are defined as the irresolute judgment aggregation rules in Definition 2.

Let $\mathcal{A}_g$ be the agenda corresponding to a goal g considered by a group of agents N and let π_g be the profile of the members judgment regarding $\mathcal{A}_g$. We define the group attitudes regarding a goal g, *i.e.*, the *decision*, to be the collective judgment set of the group.

Definition 86 (Decision). *Given a profile π_g for a considered goal g and a judgment aggregation rule F, the group N's decision regarding g is $\mathfrak{D}_g = \{A_N a \mid a \in f(\pi)\}$.*

Proposition 6.2.4. *Every group member accepts the group decision.*

Proof. As a direct consequence of axiom (**PAccess**), when the group has intention $I_N g$, every agent in N accepts that this is the group's intention, regardless of what their individually accepted regarding Gg. Also, as a consequence of axiom (**NAccess**), when the group rejects a goal, $A_N \neg Gg$, every agent i accepts this group decision. The same holds for the group beliefs. □

6.2.4 Judgment aggregation rules for agreeing on intentions

Which of the rules we introduced in Chapters 2 and 3 are adequate for use by a group that needs to agree on its intentions? Since the judgments in the intention agreement problem are binary and unweighted, we can choose both from the rules of Chapter 2 and Chapter 3. To select the particular one we need to look at the properties, defined in Chapter 4, that the rule should satisfy.

The judgment aggregation rule we can use for obtaining group goals should produce decisions that are complete and it should satisfy collective rationality. If $F(\pi)$ is not complete it is difficult to revise the group intentions. For example, if the decision contains only a group goal acceptance, then we do not know why the goal was (not) adopted and consequently when to revise it. For example, the cleaning crew decides for the goal g_3 (to collect recyclables), without having the reasons like p_9 (a container where to put them). If the information about the world is updated and $\neg p_9$ holds, the robots will continue to collect recyclables. If the aggregation of an admissible profile is not consistent with the constraints, we would not be generating reasons for the group goal. All the rules we proposed satisfy collective rationality. The rules from Chapter 2 do not produce complete judgment sets. In this case, the judgment sets can be completed by adding the missing judgments in such a way that the consistency of the set is not perturbed. For example, consider the pre-agenda $\mathcal{A} = \{p, q, p \rightarrow q\}$ and assume that the collective judgment set is $\{p, \neg q\}$. This set can be made complete by extending it into $\{p, \neg q, \neg(\mathbf{p} \rightarrow \mathbf{q})\}$.

Since the judgment aggregation problem is that of producing an agreement between the agents, the rules should be such to guarantee that the decision is responsive to the acceptances of the individuals. This means that the decision should be supported by the majority whenever that is feasible. Therefore the rule applied should be majority-preserving.

There are cases, when the profile is not majority consistent, when the rule leads the group to adopting a goal that neither of the agents endorses individually. To avoid this, we need to aggregate using a rule that satisfies the unanimity principle, if only the weak form of the property, particularly on issue Gg. As we can observe in Table 4.14, the rules that satisfy weak unanimity are R_Y, R_{MSA}, R_{RA} and R_{RY}.

Another desirable property for consensual group's decision-making contexts is resoluteness. Since all judgment aggregation rule are by construction irresolute, one can prefer the rules that produce less judgment set. It can happen that R_{MSA} generates more judgment sets than R_{RA} for the same profile. As it can be observed in Table 8.1.2, if a collective judgment is in all the judgment sets obtained by R_{RA}, then it will be in all the judgment sets obtained by R_{MSA}, for the same profile, but the reverse does not hold. This means that, for the same profile, there are judgment sets generated by R_{MSA} but not by R_{RA}. applied to a profile produces strictly more judgment sets then R_Y applied to the same profile. Therefore we exclude R_{MSA} from the set of choices.

The rule R_{RA} can be seen as better then R_Y and R_{RY} since it satisfies one more property, the strict insensitivity to reinforcement of collective judgments. In addition, one can easily construct a linear time algorithm, with respect to the size of the profile, for calculating R_{RA}, while R_Y and R_{RY} can be expected to be computationally more complex.

The irresoluteness of the rules R_Y, R_{RA} and R_{RY} must be resolved. In Chapter 5 the group was hierarchical so the agent responsible for the decision, the initiator, chose between two possible judgment set. In this context, a tie-breaking mechanism needs to be specified. Ties can be broken by randomly selecting one of the collective judgment sets, or by selecting the judgment set that contains a judgment on a particularly important issue, such as the goal, supported by the majority. In the case of the rule R_{RA} this last approach is not applicable, since the collective judgments are selected in order of strength of the majority that supports them.

6.3 The generation of multiple group goals

The mental state of the group is determined by the mental states of the members and the choice of judgment aggregation function. We represent the mental state of the group by a set Υ of AGE_{LTL} formulas. The set Υ contains the set of all candidate goals $\mathcal{G} \subseteq \mathcal{L}_G/\mathcal{L}_{prop}$ for the group and, for each $Gg \in \mathcal{G}$, the corresponding constraints $\mathcal{R}_g$, as well as the individual and collective acceptances made in the group regarding agenda $\mathcal{A}_g$. The set Υ is common knowledge for the group members. An agent uses Υ when it acts as a group member and its own beliefs and goals when it acts as an individual.

To deal with multiple, possibly mutually inconsistent goals, the group has a priority order

$\gtrsim_x$ over the group goals $\mathcal{G} \subset \Upsilon$. To avoid overburdening the language with a $\gtrsim_x$ operator, we incorporate the priority order within the constraints $\mathcal{R}^{just}_{g_i} \equiv \Gamma_i \leftrightarrow Gg_i$. We want the constraints to capture that if Gg_i is not consistent (according to the coordination rules) with some higher priority goals $Gg_1, \ldots, Gg_m$, then the group can accept Gg_i if and only if none of $Gg_1, \ldots, Gg_m$ is accepted. Hence, we replace the justification rule $\mathcal{R}^{just}_{g_i} \in \Upsilon$ with $\mathcal{R}^{pjust}_{g_i} \equiv (\Gamma_i \wedge \bigwedge_j^m (A_N \neg Gg_j)) \leftrightarrow Gg_i$, where $Gg_j \in \mathcal{G}$, $Gg_j \gtrsim_x Gg_i$ and $Gg_i \wedge Gg_j \wedge \mathcal{R}^{coord}_{g_i} \models \bot$.

Example 6.3.1. *Consider the goals and rules of the robot crew C from Example 6.2.3. Assume the crew has been given the priority order $Gg_1 >_\Upsilon Gg_2 >_\Upsilon Gg_3$. Υ contains: $\mathcal{G} = \{Gg_1, Gg_2, Gg_3\}$, one background knowledge rule, one coordination rule, three justification rules, out of which two are new priority modified rules:*
$\{\mathcal{G}, \neg p_4 \rightarrow \neg p_5, (Gg_2 \wedge \neg(Gg_1 \vee Gg_3)) \vee \neg Gg_2,\ Gg_1 \leftrightarrow (p_1 \wedge (p_2 \vee p_3)),$
$Gg_2 \leftrightarrow (p_4 \wedge p_5 \wedge (p_6 \vee p_7) \wedge A_C \neg Gg_1), Gg_3 \leftrightarrow (p_8 \wedge p_9 \wedge p_3 \wedge (A_C \neg Gg_2)\}.$

The agents give their judgments on one agenda after another starting with the agenda for the highest priority candidate goal. Once the profile π and the decision $\mathfrak{D}_g$ for a goal g are obtained, they are added to Υ. To avoid the situation in which the group casts judgments on an issue that has already been decided, we need to remove decided issues from $\mathcal{A}_g$ before eliciting the profile for this agenda.

The group goals are generated by executing **GenerateGoals**(Υ, N).

`function` **GenerateGoals**(Υ, S):
`for each` $Gg_i \in \mathcal{G}$ s.t. $[\forall Gg_j \in \mathcal{G}$: $(Gg_j \gtrsim Gg_i) \Rightarrow (A_N Gg_j \in \Upsilon$ or $A_N \neg Gg_j \in \Upsilon)]$
{ $B := (\{a \mid A_N a \in \Upsilon\} \cup \{\neg a \mid A_N \neg a \in \Upsilon\}) \cap \mathcal{A}_{g_i}$;
$\mathcal{A}^*_{g_i} := \mathcal{A}_{g_i}/B$;
$\pi_{g_i} := elicit(S, \mathcal{A}^*_{g_i}, \Upsilon)$;
$\Upsilon := \Upsilon \cup \pi_{g_i} \cup f^a(\pi_{g_i})$; }
`return` Υ.

GenerateGoals does not violate the candidate goal preference order and it terminates if *elicit* terminates. *elicit* requests the agents to submit complete judgment sets for $\pi_{g_i} \subset \Upsilon$. We require that *elicit* is such that for all returned π it holds: $\Upsilon \cup f(\pi)) \not\models \bot$ and $\Upsilon \cup \pi \triangleright i \not\models \bot$ for every $i \in N$. When a higher priority goal Gg_i is accepted by the group, a lower priority incompatible goal Gg_j cannot be adopted regardless of the judgments on the issues in $\mathcal{A}_{g_j}$. Nevertheless, *elicit* will provide individual judgments for the agenda $\mathcal{A}_{g_j}$. If the acceptance of Gg_i is reconsidered, we can obtain a new decision on Gg_j because the profile for Gg_j is available.

Example 6.3.2. *Consider the Υ sets for the robots given in Example 6.3.1. The following calls to elicit are made in the given order. First, $\pi_{g_1} = elicit(N, \mathcal{A}^*_{g_1}, \Upsilon)$ with the $GenerateGoals(\Upsilon) = \Upsilon' = \Upsilon \cup \pi_{g_1} \cup f^a(\pi_{g_1})$. Second, $\pi_{g_2} = elicit(N, \mathcal{A}^*_{g_2}, \Upsilon')$, with $GenerateGoals(\Upsilon') = \Upsilon'' = \Upsilon' \cup \pi_{g_2} \cup f^a(\pi(g_2))$. Last, $\pi_{g_3} = elicit(N, \mathcal{A}^*_{g_3}, \Upsilon'')$, with $GenerateGoals(\Upsilon'') = \Upsilon''' = \Upsilon'' \cup \pi_{g_3} \cup f^a(\pi_{g_3})$. Since there is no overlapping between agendas $\mathcal{A}_{g_2}$ and $\mathcal{A}_{g_1}$, $\mathcal{A}^*_{g_1} \equiv \mathcal{A}_{g_1}$ and $\mathcal{A}^*_{g_2} \equiv \mathcal{A}_{g_2}$. However, since $\mathcal{A}_{g_2} \cap \mathcal{A}_{g_3} = p_3$, then $\mathcal{A}^*_{g_3} = \{p_8, p_9, Gg_3\}$.*

6.4 Commitment strategies

The group can choose to reconsider the group goal in presence of new information – "a joint commitment must be *terminated* jointly" (Gilbert, 2007, pg. 143). Whether the group chooses to reconsider depends on how committed it is to the group intention corresponding to that goal. We defined the group intention to be $I_N g \equiv A_N Gg$, *i.e.* the decision to accept g as the group goal. The level of persistence of a group in their collective decision depends on the choice of commitment strategy.

These are the three main commitment strategies (introduced by (Rao and Georgeff, 1993)):
Blind commitment: $I_i g \rightarrow (I_i g \mathbf{U} B_i g)$
Single-minded commitment: $I_i g \rightarrow (I_i g \mathbf{U}(B_i g \vee B_i \Box \neg g))$
Open-minded commitment: $I_i g \rightarrow (I_i g \mathbf{U}(B_i g \vee \neg G_i g))$

These commitment strategies only consider the relation between the intention and the beliefs regarding g and Gg. In our model of group intentions, a commitment is to a goal acceptance. This enables intention reconsideration upon new information on either one of the agenda issues in $\mathcal{A}_g$, as well as on a higher priority goal.

The strength of our framework is exhibited in its ability to describe the groups' commitment not only to its decision to adopt a goal, but also to its decision to reject a goal. Namely, if the agents decided $I_N g_i$ and $A_N \neg Gg_j$, they are committed to both $I_N g_i$ and $A_N \neg Gg_j$. Commitment to reject g allows for g to be reconsidered and eventually adopted if the state of the world changes.

Let N be a set of agents with a set of candidate goals $\mathcal{G}$. Let $Gg_i, Gg_j \in \mathcal{G}$ have agendas $\mathcal{A}_{g_i}$, $\mathcal{A}_{g_j}$. We use $p \in \overline{\mathcal{A}}^p_{g_i}$ and $q_i \in \overline{\mathcal{A}}^c_{g_i}$, $q_j \in \overline{\mathcal{A}}^c_{g_j}$. The profiles and decisions are π_{g_i} and $f(\pi_{g_i})$; $Gg_j > Gg_i$, and Gg_j cannot be pursued at the same time as Gg_i.

We use the formulas $(\alpha_1) - (\alpha_5)$ to refine the blind, single-minded and open-minded commitment. Instead of the *until*, we use the temporal operator *release*: $\psi \mathbf{R} \varphi \equiv \neg(\neg\psi \mathbf{U} \neg\varphi)$, meaning that φ has to be true until and including the point where ψ first becomes true; if ψ never becomes true, φ must remain true forever. Unlike the *until* operator, the *release* operator does not guarantee that the right hand-side formula will ever become true, which in our case translates to the fact that an agent could be forever committed to a goal.

(α_1) $Eg_i \mathbf{R} I_N g_i$
(α_2) $\bot \mathbf{R} A_N \neg Gg_i$
(α_3) $(E\Box\neg g_i \vee Eg_i)\mathbf{R} A_N q_i$
(α_4) $A_N \neg q_j \mathbf{R} A_N q_i$
(α_5) $A_N p \rightarrow (E\neg p \mathbf{R} A_N q_i)$

Blind commitment: $\alpha_1 \wedge \alpha_2$.
Only the observation that the goal is achieved (Eg_i) can release the intention to achieve the goal $I_N g_i$. If the goal is never achieved, the group is always committed to it. If a goal is not accepted, then the agents do not reconsider accepting it.

Single-minded commitment: α_3.
Only new information on the goal (either that the goal is achieved or had become impossible) can release the decision of the group to adopt /reject the goal. Hence, new information is only regarded if it concerns the conclusion, while information on the remaining agenda items is ignored.

Extended single-minded commitment: $\alpha_3 \wedge \alpha_4$.
Not only new information on Gg_i, but also the collective acceptance to adopt a more important incompatible goal Gg_j can release the intention of the group to achieve Gg_i. Similarly, if

Gg_i is not accepted, the non-acceptance can be revised, not only if Gg_j is observed to be impossible or achieved, but also when the commitment to pursue Gg_j is dropped (for whatever reason).

Open-minded commitment: $\alpha_3 \wedge \alpha_5$.
A group maintains its collective acceptances to adopt or reject a goal as long as the new information regarding all collectively accepted agenda items is consistent with $f(\pi_{g_i})$.

Extended open-minded commitment: $\alpha_3 \wedge \alpha_4 \wedge \alpha_5$.
Extending on the single-minded commitment, a change in intention to pursue a higher priority goal Gg_j can also release the acceptance of the group on Gg_i.

Once an intention is dropped, a group may need to reconsider its collective acceptances. This may cause for the dropped goal to be re-affirmed, but a reconsideration process will be invoked nevertheless.

6.5 Reconsideration of group attitudes

In Section 6.3 we defined the mental state of the group Υ. We can now define what it means for a group to be *coherent*.

Definition 87 (Group coherence)**.** *Given a Kripke structure $\mathcal{M}$ and situations $s \in W$, a group of N agents is coherent if the following conditions are met:*
(ρ_1): $\mathcal{M} \models \neg(A_S a \wedge A_S \neg a)$ for any $S \subseteq N$ and any $a \in \mathcal{A}_g$.
(ρ_2): If $\mathcal{M}, s \models \Upsilon$ then $\Upsilon \not\models \bot$.
(ρ_3): $\mathcal{M}, s \models \bigwedge \mathcal{G} \rightarrow \neg\Box\neg g$ for all $Gg \in \mathcal{G}$.
(ρ_4): Let $Gg \in \mathcal{G}$ and $\mathcal{G}' = \mathcal{G}/\{Gg\}$, then $\mathcal{M} \models (\bigwedge \mathcal{G} \wedge E\Box\neg g) \rightarrow \boldsymbol{X}(\neg Gg)$.
(ρ_5): Let $p \in \mathcal{A}_g^p$ and $q \in \{Gg, \neg Gg\}$. $Ep \wedge (Ep\mathbf{R}A_N q) \rightarrow \boldsymbol{X}A_N p$

The first condition ensures that no contradictory judgments are given. The second condition ensures that the mental state of the group is logically consistent in all situations. The third and fourth conditions ensure that impossible goals cannot be part of the set of candidate goals and if g is observed to be impossible in situation s, then it will be removed from $\mathcal{G}$ in the next situation. ρ_5 enforces the acceptance of the new information on the group level, when the commitment strategy so allows – after a is observed and that led the group to de-commit from g, the group necessarily accepts a.

A coherent group accepts the observed new information on a premise. This may cause the collective acceptances to be inconsistent with the justification rules. Consequently, the decisions and/or the profiles in Υ need to be changed in order to ensure that ρ_1 and ρ_2 are satisfied. If, however $\Box\neg g$ or g is observed, the group reconsiders Υ by removing Gg from $\mathcal{G}$. In this case, the decisions and profiles are not changed.

For simplicity, at present we work with a world in which the agents' knowledge can only increase, namely the observed information is not a fluent. A few more conditions need to be added to the definition of group coherence for our model to be able to be applicable to fluents. For example, we need to define which observation is accepted when two subsequent contradictory observations happen.

6.5.1 Reconsideration strategies

For the group to be coherent in all situations, the acceptances regarding the group goals need to be reconsidered after de-commitment. Let $\mathfrak{D}_g \subset \Upsilon$ contain the group acceptances for a goal g, while $\pi_g \subset \Upsilon$ contain the profile for g. There are two basic ways in which a collective judgment set can be reconsidered. The first way is to elicit a new profile for g and apply judgment aggregation to it to obtain the reconsidered $\mathfrak{D}_g^*$. The second is to reconsider only $\mathfrak{D}_g$ without re-eliciting individual judgments. The first approach requires communication among agents. The second approach can be done by each agent reconsidering Υ by herself. We identify three reconsideration strategies available to the agents. The strategies are ordered from the least to the most demanding in terms of agent communication.

Decision reconsideration ($\mathcal{D}$-r).

Assume that Ea, $a \in \overline{\mathcal{A}}_g^p$, $q \in \{Gg, \neg Gg\}$ and the group de-committed from $A_N q$. The reconsidered decision $\mathfrak{D}_g^*$ is such that a is accepted, i.e., $A_N a \in \mathfrak{D}_g^*$, and the entire decision is consistent with the justification rules, namely $\mathbf{R}u_g^{pjust} \cup \mathfrak{D}_g^* \not\models \bot$. If the $\mathcal{D}$-r specifies a unique $\mathfrak{D}_g^*$, for any observed information and any $\mathfrak{D}_g$, then Υ can be reconsidered without any communication among the agents. Given the form of $\mathcal{R}_g^{pjust}$ (see Section 6.3), this will always be the case.

However $\mathcal{D}$-r is not always an option when the de-commitment occurred due to a change in collective acceptance of a higher priority goal g'. Let $q' \in \{Gg', \neg Gg'\}$. Let the new acceptance be $A_N \neg q'$. $\mathcal{D}$-r is possible if and only if $\mathfrak{D}_g^* = \mathfrak{D}_g$ and $\mathcal{R}_g^{pjust} \cup \mathfrak{D}_g \cup \{A_N \neg q'\} \not\models \bot$. Recall that $A_N q'$ was not in $\mathcal{A}_g$ and as such the acceptance of q' or $\neg q'$ is never in the decision for π_g.

Partial reconsideration of the profile (Partial π-r).

Assume that Ea, $a \in \overline{\mathcal{A}}_g$, $Gg \in \mathcal{G}$. Not only the group, but also the individual agents need to accept a. The *Partial π-r* asks for new individual judgments to be elicited. This is done to ensure the logical consistency of the individual judgment sets with the observations. New judgments are only elicited from the agents i which $A_{\{i\}} \neg a$.

Let $W \subseteq N$ be the subset of agents i s.t. $A_{\{i\}} \neg a \in \Upsilon$. Agents i are s.t. $A_{\{i\}} a \in \Upsilon$ when the observation is $E\neg a$. Let $\pi_g^W \subseteq \pi_g$ be the set of all acceptances made by the agents in W. We construct $\Upsilon' = \Upsilon / \pi_g^W$. The new profile and decision are obtained by executing *GenerateGoals* (Υ', W).

Example 6.5.1. *Consider Example 6.2.3. Assume that* $\mathfrak{D}_{g_1} = \{A_C p_1, A_C \neg p_2, A_C p_3, A_C Gg_1\}$, $\mathfrak{D}_{g_2} = \{A_C p_4, A_C p_5, A_C p_6, A_C p_7, A_C \neg Gg_2\}$ *and* $\mathfrak{D}_{g_3} = \{A_C p_8, A_C p_9, A_C Gg_3\}$ *are the group's decisions. Assume the group de-commits on* Gg_1 *because of* $E\neg p_2$*. If the group is committed to* Gg_3*, the commitment on* Gg_3 *will not allow for* $A_N p_3$ *to be modified when reconsidering* Gg_1*. Since* $A_N p_3$ *exists in* Υ'*,* p_3 *will be excluded from the (new) agenda for* g_1*, although it was originally in it. elicit calls only on the agents in W to complete* $\pi_{g_1} \in \Upsilon'$ *with their judgment sets.*

Full profile reconsideration (Full π-r).

The full profile reconsideration is the same as the partial reconsiderations except now $W = N$. Namely, within the full profile revision strategy, each agent is asked to revise his judgment set by accepting the new information, regardless of whether he had already accepted it.

6.5.2 Combining revision and commitment strategies

Unlike the commitment strategies of (Rao and Georgeff, 1993), in our framework the commitment strategies are not axioms of the logic. We require that the commitment strategy is valid in all the models of the group and not in all the models of AGE_{LTL}. This allows the group to define different commitment strategies and different revision strategies for different goals. It might even choose to revise differently depending on which information triggered the revision. Choosing different revision strategies for each goal, or each type of new information, should not undermine the coherence of the mental state of the group, the set Υ. The conditions of group coherence of the group ensures that, after every reconsideration, Υ must remain consistent. However, some combinations of commitment strategies can lead to incoherence of Υ.

Example 6.5.2. *Consider the decisions in Example 6.5.1 . Assume that initially the group is such that it follows the open-minded commitment for $I_C g_1$ and blind commitment for $I_C g_3$, with goal open-minded commitment for $A_C \neg G g_2$. If $E g_1$ and thus $I_C g_1$ is dropped, then the extended open-minded commitment would allow $A_C \neg G g_2$ to be reconsidered and eventually $I_C g_2$ established. However, since the group is blindly committed to $I_C g_3$, this change will not cause reconsideration and as a result both $I_C g_2$ and $I_C g_3$ will be in Υ, thus making Υ incoherent.*

Problems arise when $sub(\mathcal{R}_{g_i}^{pjust}) \cap sub(\mathcal{R}_{g_j}^{pjust}) \neq \emptyset$, where $sub(\mathcal{R}_{g}^{pjust})$ denotes the set of atomic sub-formulas of some goal g and $Gg_i, Gg_j \in \mathcal{G}$. Proposition 6.5.3 summarizes under which conditions these problems are avoided.

Proposition 6.5.3. *Let α' and α'' be the commitment strategies selected for g_i and g_j correspondingly. $\Upsilon \cup \alpha' \cup \alpha'' \not\models \bot$ (in all situations):*
a) if $\varphi \in sub(\mathcal{R}_{g_i}^{pjust}) \cap sub(\mathcal{R}_{g_j}^{pjust})$ and $p \in \mathcal{A}_{g_i} \cap \mathcal{A}_{g_j}$, then α_5 is either in both α' and α'' or in none;
b) if Gg_i is more important than Gg_j while Gg_j and Gg_i cannot be accepted at the same time, then $\alpha_4 \in \alpha''$.

Proof. The proof is straightforward. If the change in the group (non)acceptance of Gg_i causes the $A_N Gg_j$ to induce group incoherence, then we are able to de-commit from $A_N Gg_j$. If we would not able to de-comit from $A_N Gg_j$ then group coherence would be blocked. If the change in the group (non)acceptance of Gg_i is caused by an observation on a premise $p \in \mathcal{A}_{g_i} \cap \mathcal{A}_{g_j}$ then condition a) ensures that the commitment to $A_N Gg_j$ does not block group coherence. If the change on $A_N Gg_j$ is caused by a change in commitment to a higher priority goal, the condition b) ensures that a commitment regarding Gg_j does not block group coherence. Condition b) allows only "goal sensitive" commitments to be selected for lower level goals. □

Commitment to	Release on				Change	How		
$A_N(\neg)Gg$	$\Box\neg g$	g	Gg_j	$\mathcal{A}_g^p$	Υ	$\circledast\mathfrak{D}_g$	$\circledast\pi_g$	JA
Blind	✓							
Single-minded	✓	✓			$\mathfrak{D}$-r	✓		
Extended	✓	✓	✓		Partial π-r		✓	✓
Open-minded	✓	✓		✓	Full π-r		✓	✓
Extended	✓	✓	✓	✓				

Table 6.1: $Gg_j > Gg$ and cannot be pursued at the same time with Gg. $\circledast\mathfrak{D}_g$ denotes collective attitudes for g are reconsidered. $\circledast\, \pi_g$ denotes the profile (all or some parts of it) is re-elicited.

6.6 Conclusions

We present a formalization of non-summative beliefs, goals and intentions in AGE_{LTL} and show how they can be generated using judgment aggregation. Our multi-agent AGE_{LTL} logic extends BDI_{LTL}. In accordance with the non-summative view, having a group intention $I_N g$ in our framework does not imply $I_{\{i\}}g$ for each the member i. We extended the commitment strategies of (Rao and Georgeff, 1993) to increase the reactivity of the group to new information. Now the commitment strategies are not axioms of the representation logic; instead they are a property of a group. Groups can have different levels of commitment to different goals. We showed how the group can combine different commitments to different goals.

An advantage of our framework is its ability to allow groups to commit to a decision to reject a goal, thus having the option to reconsider rejected goals. Furthermore, we do not only show when to reconsider, but also how, by defining reconsideration strategies. Table 1 summarizes our commitment and reconsideration strategies.

In our framework, the entire group observes the new information. One can also explore the case when only some members of the group observe the new information. The only assumptions we make regarding the connectivity of the members is that they are able to communicate their acceptances and receive the aggregation result. The problem of elicitation and communication complexity in voting is a nontrivial one (Conitzer and Sandholm, 2002b, 2005) and in the future we intend to study these properties of our framework.

In our framework, the group has an intention g if it has agreed to pursue g as a group goal. The agents agree on which goal to pursue by stating their acceptances of a proposed group goal, and related beliefs that support or justify that goal, and applying a judgment aggregation rule on these acceptances. Since the judgment aggregation rule serves as an agreement reaching mechanism, it needs to satisfy the properties of majority-preservation and the unanimity principle. Based on these properties we propose that the rules R_{RA}, R_Y and R_{RY}, developed in Chapter 2 are used.

The rules from Chapter 2 are defined for judgment aggregation problems that are specified by an agenda and a set of agent names. A set of constraints is not part of the problem specification. We use the set of constraints here to describe the added knowledge that the group needs to take into consideration when choosing its goals. We are able to use the rules from Chapter 2 anyway since we defined an admissibility criterion for the individual judgment sets.

Although we studied the complexity of winner determination for the rules of Chapter 3 we did not make the same analysis for the rules of Chapter 2. To choose further between R_{RA}, R_Y and R_{RY} we need to know their complexity-theoretic properties as well.

In this context of agreeing on group goals we imposed the requirement that the agents are always able to declare whether they accept a belief or a group goal or not. This requirement was feasible because the agent's judgments are expressions of acceptance, not an estimate of the state of the world as in Chapter 5. Nevertheless, it might be desirable to give the option to an agent, or a group, to also express express "do not know whether to accept a" via the formula $\neg A_{\{i\}} a \wedge \neg A_{\{i\}} \neg a$, or respectively $\neg A_N a \wedge \neg A_N \neg a$. This option implies a change in the input value-type of the judgment aggregation problem from binary to ternary. To be able to allow for this option we need to consider the extension of the rules of Chapter 2 to ternary judgments.

In the work we presented, we do not consider how an individual constructs his judgments. We can take that $B_i\, \varphi \rightarrow A_{\{i\}} \varphi$, but this is not a requirement for all agents. We would expect "honest" agents to follow this rule, but we can also define dishonest agents for which $B_i \varphi \rightarrow A_{\{i\}} \varphi$ does not hold. In the latter case, the agent might declare $A_{\{i\}} \varphi$ while it does not believe φ. Given that the group attitudes are established by an aggregation rule that can be expected to be, as almost all but the most trivial social choice rules, manipulable, the question is whether there are scenarios in which an agent can have the incentive to behave strategically in rendering judgments. Furthermore, given that some of the reconsideration strategies call for re-elicitation of judgments, can an agent have the incentive to behave strategically in rendering judgments that would lead to sooner re-elicitation? To answer these questions we need to study the manipulability properties of our rules.

7

Related work

The problem of aggregating profiles of yes/no decisions over a set of logically related issues is, with minor variations, studied under different names: judgment aggregation, majority-voting under interconnected decisions, abstract binary aggregation, aggregation of binary evaluations. Some comparisons of the differences between the frameworks denoted with the different terms can be found in (List and Polak, 2010; Grandi and Endriss, 2010). We uniformly use the term judgment aggregation, although strictly speaking, our rules based on minimization are judgment aggregation rules and our distance-based rules are both judgment aggregation and ternary-evaluation aggregation rules.

To the best of our knowledge there is no work that is related to the thesis as a whole. Therefore we give the related work with respect to our judgment aggregation rules, properties and models of collective reasoning.

7.1 Part I- designing rules

The study of judgment aggregation rules *per se* is a rarity in judgment aggregation. The only exception in this sense is probably the work of (Miller and Osherson, 2009). Some judgment aggregation rules do appear in the context of other works.

7.1.1 Rules based on minimization

The premise-based procedure has been introduced in (Kornhauser and Sager, 1993) under the name "issue-by-issue voting" and studied extensively in (Dietrich and Mongin, 2010). The conclusion-based procedure has been studied in (Pigozzi et al., 2009). These were the first rules considered in the literature. The sequential procedures have been introduced in (List, 2004a) studied also in (Dietrich and List, 2007b; Li, 2010), followed by quota-based rules (Dietrich and List, 2007b; Dietrich, 2010). The quota-based rules are a class of rules where each proposition of the agenda is associated with a quota, and the proposition is accepted only if the proportion of individuals accepting it is above the quota. The majority rule is a special case of quota-based rules. Lastly, the distance-based rules are studied in (Miller and Osherson, 2009) and in (Pigozzi, 2006). We explain the principles of these rules in Chapter 2, therefore here we just enlist the relations between these rules and the rules we propose.

The rules R_{MSA}, R_{MWA} and R_{MNAC} are special cases of three of the four distance-based rules introduced in (Miller and Osherson, 2009). More precisely, for d being the Hamming distance, R_{MSA} is theory equivalent with $Endpoint_d$; R_{MWA} and R_{MNAC} are equivalent with $Prototype_d$ and $Full_d$ correspondingly.

Nehring et al. (2011) and (Nehring and Pivato, 2011) define the *Condorcet admissible set* to be any maximally consistent sub-set of the majoritarian set. Recall that the majoritarian set for a profile is the set of all judgments supported by a majority of agents according to the profile. The Condorcet admissible sets for any profile are equivalent to the set of judgment sets that can be produced by applying R_{MSA} to that profile. Rather then considering a rule that derives the Condorcet admissible sets, they consider it a space of judgment sets and study under which conditions this space collapses to a singleton for a profile. They call *Condorcet determinate* what we define as majority-consistent profile.

Nehring et al. (2011) introduce two new judgment aggregation rules: the *Slater rule* and the *Median rule*. The Slater rule is equivalent to our R_{MCSA} and the Median rule to our R_{MWA}. Nehring and Pivato (2011) in addition introduce the *LexiMin* rule which we defined as R_{RA}. The difference between R_{RA} and R_{MWA} on one side and the LexiMin and Median rule on the other, is that the LexiMin and Median rules are defined for agent-weighted judgments. Nehring et al. (2011) and also (Nehring and Pivato, 2011) consider agent weights, normalized over $[0,1]$ so that the sum of the weights of all agents on an issue is always 1. The focus of (Nehring et al., 2011) is to characterize the conditions, and identify the likelihood, under which $T_{R_{MSA}}(P)$ contains a complete judgment set, some particular element of the agenda or the full set of consistent and complete judgment sets.

7.1.2 Distance-based non-binary rules

Distance-based rules for aggregating judgments have been inspired by the *model based* distance-based belief merging rules (Konieczny et al., 2004). The first distance-based rules for aggregating judgments have been introduced by (Pigozzi, 2006) who observed that there are considerable similarities between the belief merging and the judgment aggregation problems.

Belief merging theory studies the problem of merging belief bases. Given a set of belief bases and a set of constraints *IC*, the problem of merging the belief bases is to generate a belief base which satisfies all of the *IC* constraints and incorporates a maximal amount of beliefs from the bases that are merged. In judgment aggregation the individual judgment sets are cast on agenda issues. In belief merging no agenda is defined.

In belief merging, the primary concern of the merging process is to maximize the information content from the merged belief bases. Rules for merging beliefs are constructed so that they satisfy a set of postulates that are inspired from the belief revision postulates (Alchourrón et al., 1985). Judgment aggregation rules are constructed so that they satisfy a set of properties that are inspired from voting theory and preference aggregation (List and Polak, 2010). The aggregation properties are not concerned with maximizing the information content from the individual judgment sets.

Pigozzi (2006) applies the model based operator of (Konieczny et al., 2004), defined for a Hamming distance and $\sum$, directly by treating the judgment sets $\hat{A}_i$ as belief, or knowledge bases. The agenda is defined as a set of propositional logic *atoms*. A judgment set is a consistent set of atomic and non-atomic formulas, complete for the agenda $\mathcal{A}$. The set of constraints *IC* corresponds to the set of rules $\mathcal{R}$. Each judgment set is consistent with respect to $\mathcal{R}$. The set of all models for $\mathcal{R}$ is the co-domain of the aggregation rule. Endriss et al. (2010b) also define their procedure in terms of the Hamming distance and $\sum$, but contrary to

(Pigozzi, 2006), they define the agenda in such a way that it can contain non-atomic formulas. However they do not consider additional rules and in their framework $\mathcal{R} = \emptyset$. Consequently, the distance-based rule for the Hamming distance and the $\sum$ as an aggregation operator are not equivalent. The distance-based rules we propose allow for both non-atomic formulas in the agenda and additional rules to be externally specified.

The use of agent weights in distance-based belief merging contexts has been previously be considered in (Revesz, 1995). Weights in the context of judgment aggregation have not been considered. The weights assigned to issues, or to (agent,issue) pairs, have not been considered in belief merging or in judgment aggregation context. Distances between sequences that contain more values than binary have been considered in the literature (Condotta et al., 2008; Coste-Marquis et al., 2007).

There are many reasons for which one would like to relax the requirement of completeness for judgment sets and allow the agents to abstain on some issues. That the requirement of completeness for judgment sets is too strict, has been discusses by (Gärdenfors, 2006). Dokow and Holzman (2010b) construct a framework for binary aggregation in which abstentions are allowed. The value assigned to the abstention is descriptive, namely a special symbol "$*$", while 0 and 1 are used for the "no" and "yes" judgments correspondingly. Pauly and van Hees (2006) construct a framework for multivalued logics in which the judgment sets with abstentions may be seen as a special case. Dietrich (2007) constructs a general logic framework in which some ternary logics are a special case. These authors focus on proving impossibility results for their respective frameworks and offer no particular rules for aggregating judgment sequences with abstentions.

One might argue that the aggregation of judgments with abstentions poses no particular challenge; namely, if an agent abstains then his input can be ignored, as if a judgment was not elicited from that agent. This perceived simplicity disappears once one is reminded that the agenda issues are logically related and the judgments assigned to them are logically related. Abstaining on one-issue influences the judgments that can be assigned to the rest of the issues, by the same agent and on the generation of collective judgments for all issues. The challenge in developing judgment aggregation rules that handle abstentions is not in the representation of the abstentions but in the interpretation of the abstentions. When distance-based rules are used, the interpretation assigned to abstentions hinges upon the set $\mathbb{A}$ and the chosen distance metric. The set $\mathbb{A}$ depends on the particular ternary logic used for representing the judgments.

The distance-based rules we introduced can be applied in the framework of (Pauly and van Hees, 2006) when $\mathbb{A}$ is constructed using the logic of (Post, 1921). The rules can be applied in the framework of (Dokow and Holzman, 2010b), when $*$ is treated as $\frac{1}{2}$ and either the logic of (Łukasiewicz, 1920) or the logic of (Kleene, 1938) is used. The impossibility results proved by (Dietrich, 2007) hold for judgment aggregation problems represented by the Łukasiewicz logic since this logic is a special case of the general logic introduced by (Dietrich, 2007). We show that this holds.

Dietrich (2007) defines the properties **L1** $-$ **L3** for $(\mathcal{L}, \models)$ and proves his impossibility of aggregation results for general logics that satisfy **L1** $-$ **L3** (Dietrich, 2007, pg. 554).

L1 For any $p \in \mathcal{L}$ it holds that $p \models p$ (*self-entailment*).

L2 For any $p \in \mathcal{L}$ and $S_1 \subseteq S_2 \subseteq \mathcal{L}$, if $S_1 \models p$ then $S_2 \models p$ (*monotonicity*).

L3 The empty set $\emptyset$ is consistent, and each consistent set $S_1 \subseteq \mathcal{L}$ has a consistent super-set

$S_2 \subseteq \mathcal{L}$ containing a member of each pair $p, \neg p \in \mathcal{L}$ (*completability*).

Proposition 7.1.1. *The pair* $(\mathcal{L}_L, \models_3)$*, where* $\mathcal{L}_L$ *is the logic of Łukasiewicz, satisfies self-entailment* **L1***, monotonicity* **L2***, and completability* **L2***.*

Proof. Self-entailment follows directly from the definition of the operator $\models_3$. Namely $\Gamma \models_3 \psi$ if and only if $\bigwedge \Gamma \rightarrow \psi$.
When $S_1 \models_3 \varphi$ then for all valuation v, if $v(\bigwedge S_1) = 1$ then $v(\varphi) = 1$ and if $v(\bigwedge S_1) = \frac{1}{2}$ then $v(\varphi) \in \{1, \frac{1}{2}\}$. We prove monotonicity by distinguishing among cases:

$v(\bigwedge S_1) = 1$	$v(\varphi) = 1$	$v(\bigwedge S_1/S_2) = 1$	$v(S_2) =$ hence $S_2 \models_3 \varphi$
$v(\bigwedge S_1) = 1$	$v(\varphi) = 1$	$v(\bigwedge S_1/S_2) = \frac{1}{2}$	$v(S_2) = \frac{1}{2}$ hence $S_2 \models_3 \varphi$
$v(\bigwedge S_1) = 1$	$v(\varphi) \in 1$	$v(\bigwedge S_1/S_2) = 0$	$v(S_2) = 0$ hence $S_2 \models_3 \varphi$
$v(\bigwedge S_1) = \frac{1}{2}$	$v(\varphi) \in \{1, \frac{1}{2}\}$	$v(\bigwedge S_1/S_2) = 1$	$v(S_2) = \frac{1}{2}$ hence $S_2 \models_3 \varphi$
$v(\bigwedge S_1) = \frac{1}{2}$	$v(\varphi) \in \{1, \frac{1}{2}\}$	$v(\bigwedge S_1/S_2) = \frac{1}{2}$	$v(S_2) = \frac{1}{2}$ hence $S_2 \models_3 \varphi$
$v(\bigwedge S_1) = \frac{1}{2}$	$v(\varphi) \in \{1, \frac{1}{2}\}$	$v(\bigwedge S_1/S_2) = 0$	$v(S_2) = 0$ hence $S_2 \models_3 \varphi$

The pair $(\mathcal{L}^l, \models_3)$ satisfies completability. The empty set is consistent since $v(\varnothing) =_{def} 1$. When $v(\varphi) = 1$ or $v(\varphi) = 0$, either φ or $\neg\varphi$ can be added to S, but not both. When S is valued to $\frac{1}{2}$, if $v(\varphi) = 1$ then $S \models \varphi$, but it is not true that $S \models \neg\varphi$, since $v(S) \neq 0$. For $v(\varphi) = 0$ then $S \models_3 \neg\varphi$, and it is not true that $S \models_3 \varphi$.

□

Apart from the basic ternary logics and the classical entailment operator, one can consider ternary paraconsistent logics, such as the ones studied in (Konieczny and Marquis, 2002), for representing judgments. The distance-based operators can be applied regardless of the logic, as long as the set $\mathbb{A}$ can be specified.

In addition to the completeness requirement, the judgment aggregation theory also stipulates that the judgment sets are consistent, with respect to the rules $\mathcal{R}$ when such rules are given. Miller (2008) considers the case when the constraints are subjective, namely each agent's judgment set $\hat{A}$ is consistent in terms of that agent's constraint $\mathcal{R}_i$, $\hat{A}_i \cup \mathcal{R}_i \not\models \perp$. Miller (2008) generalizes judgment aggregation to subjective decision situations, implying that the impossibility results studied in (Dietrich, 2007) persist without individual agreement on the set of constraints.

In addition to different constraints, one may consider different entailment operators $\models_i$ for each agent, thus having $\hat{A}_i \cup \mathcal{R}_i \not\models_i \perp$ as the required consistency property for the individual judgment sets. As long as the $\mathcal{R}$ and $\models_3$ are determined for the collective judgment sets, the $\mathbb{A}$ set can be constructed and our distance-based merging rules can be specified.

Benamara et al. (2010) consider the problem of aggregating judgment sets in which not only abstentions, but also *neutral judgments* are allowed. They define an agenda as a set of atoms corresponding to $\mathcal{A}^p$ and a singleton set of an atom d, called a *decision*, corresponding to $\mathcal{A}^c$. What we call an abstention here, is called a neutral judgment in (Benamara et al., 2010), denoted by "?" and representing the case in which an agent is undecided on an issue. The abstentions in (Benamara et al., 2010) are denoted by "X" and represent the case in which an agent deems the agenda issue "irrelevant" for the decision. Thus judgments are assignments $\mathcal{A} \mapsto \{0, 1, ?, X\}$. A decision rule is a formula $\Gamma(\mathcal{A}^p) \leftrightarrow \overline{d}$, where $\Gamma(\mathcal{A}^p)$ is a propositional logic formula built on literals $\overline{\mathcal{A}}^p$ and $\overline{d} \in \{d, \neg d\}$. The agents can accept, reject or be neutral regarding the rule, *i.e.*, , assign it a "value" from $\{0, 1, ?\}$. Benamara et al. (2010) give a judgment aggregation procedure for this framework, according to which the "judgment" on the rule is determined according to a majority rule, the collective judgments on the premises are determined by ignoring the abstentions and applying a rule, which we defined as majority m, to the rest. The collective judgment on the decision is reached as in the premise-based procedure if the majority accepts the rule and by applying the majority rule over the individual judgments for d.

We can represent the framework of (Benamara et al., 2010) into our framework in the following way. Let us denote the decision rule $\Gamma(\mathcal{A}^p) \leftrightarrow \overline{d}$ by α. The agenda $\mathcal{A}$ is constructed as $\mathcal{A} = \mathcal{A}^p \cup \mathcal{A}^c \cup \{\alpha\}$. We replace each judgment judgment ? with $\frac{1}{2}$. We construct a weight matrix using weights $w_{i,j} = 0$ when an agent i assigns X to issue a_j and $w_{i,j} = 1$ in every other case. The judgments X can be replaced with either one of $\{0, \frac{1}{2}, 1\}$, since they are assigned a weight 0. Then the premise-based procedure can be defined, considering α as one of the premises.

Li (2010) considers the sequential aggregation rules and allows for continued-valued judgments on the $[0, 1]$ interval. An agent expresses the strength of his acceptance or rejection of an issue through the continuous judgments. The judgment $v(\neg\varphi)$ of a negated issue φ assigned a judgment $v(\neg\varphi) = 1 - v(\varphi)$. The aim of Li's work is to determine decision paths that maximize the strength of each judgment.

7.2 Part II - properties of rules

Properties for judgment aggregation rules have been studied with the objective of determining which set of properties admits a judgment aggregation function or rule. Apart from the universal domain, anonymity and independence properties, monotonicity and unanimity were introduced. Grandi and Endriss (2010, 2011) instead study the connection between the properties of aggregation rules and the language in which the sets of constraints $\mathcal{R}$ are expressed.

The unanimity principle for aggregation functions was introduced in (Dietrich and List, 2008b). Nehring et al. (2011) define and study, what we define as, the strong unanimity principle. They observed that this principle is violated by the Condorcet admissible sets for most agendas, which is consistent with our observation that R_{MSA} does not satisfy this property.

Monotonicity as a property imposed on a subset of the agenda (to address manipulability issues) was introduced in (Dietrich and List, 2005), and monotonicity on a judgment set in (Dietrich and List, 2008a). Monotonicity as a property of judgment aggregation functions was defined in (List and Puppe, 2009).

Nehring and Pivato (2011) study two properties that are very similar with our Separability I and Separability II. They define separability as we define Separability II. Our Separability I corresponds to the reinforcement of (Nehring and Pivato, 2011), but it is a stronger property. The reinforcement property of (Nehring and Pivato, 2011) applies to entire judgment sets, whereas Separability I is applied to propositions. Therefore we obtain that R_{MWA} violates Separability, while the Median rule of (Nehring and Pivato, 2011) satisfies this reinforcement property.

What we define as independence of cloned agenda issues, has already been introduced in (Dietrich, 2006b) as the sensitivity to *logical agenda manipulation*. Dietrich (2006b) studies this property as a justification for imposing the independence of irrelevant information condition on judgment aggregation rules and functions.

7.3 Part III - applying rules

To the best of our knowledge this thesis is the first investigation in applying judgment aggregation for reaching group decisions in multiagent systems. We considered two types of groups, hierarchical and consensual, that give rise to two different aggregation contexts.

In the example of hierarchical group, one agent needs to aggregate the opinions of the other agents to reach a decision. The procedure we proposed for group decision-making is one that produces satisficing decisions. Satisficing is explored as a single agent approach to decision-making, but not for groups.

In the example of consensual group, we model non-summative group intentions, and propose a method for their generation, commitment strategies and revision strategies. Group intentions, how they are modeled, generated and revised is a question that has been considered since the advent of multi-agent systems.

7.3.1 Satisficing

As (Zilberstein, 1998) observes, there has been a search for useful techniques from decision making, since it is widely accepted that optimal decision-making is too computationally complex. The concept of satisficing (Simon, 1955) offers an alternative to the search for an optimal decision, however Simon does not instruct on how to construct satisficing algorithms or systems. Zilberstein (1998) argues that optimizing is an alternative to satisficing.

Satisficing is little used as an approach to group decision-making. However, it is in the case of group decision-making that the complexity of making an optimal decision becomes high. An exception is the work of (Stirling and Goodrich, 1999; Stirling and Nokleby, 2009) who develop satisficing games by constructing conditional utilities. Their utilities "take into account the interests of others as well as the self, represent an alternative to the categorical utilities of classical decision theory."(Stirling and Nokleby, 2009, pg.53).

The recognition-primed group decision-making model we construct is a model that allows for a team of agents to reach decisions that are satisfy a set of interrelated sufficient cues. This model is applicable when at least one agent in the team is familiar with the decision context since cues are determined from his experience. We expect that, how optimal is a group decision reached this way depends on the team's ability to learn from its mistakes.

7.3.2 Group intentions

Collective intentions are studied and formalized in (Dunin-Keplicz and Verbrugge, 2010, Chapter 3) and also, among others, in (Grosz and Hunsberger, 2007; Jennings, 1995; Singh, 1990). The dynamics of intentions have been considered in (van der Hoek et al., 2007; Grosz and Hunsberger, 2007).

In (Dunin-Keplicz and Verbrugge, 2010, Section 3.9) we find a detailed overview of the various formalizations of group intentions. Most of this work does not consider the dynamics of group intentions. An exception is (Grosz and Hunsberger, 2007). In (van der Hoek et al., 2007) the reconsideration of individual intentions and associated plans is considered.

Grosz and Hunsberger (2007) recognize that groups need to make group decisions with respect to many intention related issues such as how to change their intentions. In (Hunsberger, 2002) they conclude that a specification of group decision-making mechanism must include: (1) the possible inputs an agent can make into the mechanism; (2) the conditions under which agents may make those inputs; (3) rules for determining which combinations of agent inputs establish group decisions; and (4) a method for making the new decision known to all the members of the group. Although they consider examples of mechanisms such as unanimous approval, they make no connection with social choice theory.

We assume that the group has an order of importance for its candidate goals. Alternatively, the group can also agree on this order by expressing individual preferences. Uckelman and Endriss (2010) show how individual (cardinal) preferences over goals can be aggregated. Intentions and their role in deliberation for individual agents have been studied in a game theoretic framework by (Roy, 2009a,b). Icard et al. (2010) consider the joint revision of individual attitudes, with the revision of beliefs triggering intention revision. We allow for both the changes in epistemic and in motivational attitudes to be the cause for reconsideration.

8

Summary

People agree on things all the time. They discuss about the issues at hand and make collectively binding decisions. Some of these decisions alter human history, others alter dinner plans. Regardless of the context of the agreement, social choice rules have been developed to serve humanity when it needs to reach consents.

Since its inception, computing continues to grow more and more powerful, but at the same time more and more distributed. As a consequence, computers, processors and users, more precisely artificial agents acting on their behalf, need to reach collectively binding decisions. In this thesis we show that this problem of reaching collectively binding decisions can be solved by "computationalizing" social choice, in particular the social choice discipline of judgment aggregation.

We consider two types of groups: consensual and hierarchic. In the first, the agents reach a decision collectively, while in the second there is one agent that makes a decision by considering the input from others. A consensual group is a representation of a distributed system of agents that need to behave as a whole and make decisions that govern their actions and behavior. A hierarchical group is a representation of an agent decision-maker that needs to use, not only his own, but the knowledge, opinions and expertise of many other, possibly distributed agents. These decisions can be used by the agent himself, by the agents who contribute information or by a wider set of agents within the scope of one institution. For each of these two types of groups we give an example of a group decision-reaching problem and show how it can be solved using judgment aggregation. In both of the examples judgment aggregation is a consent reaching method, applicable even when it cannot be assumed that the agents persuade each other on a single position.

8.1 Results

8.1.1 Designing judgment aggregation rules

Judgment aggregation theory is a new discipline of social choice in the scope of which not enough effort has been devoted to constructing and analyzing specific aggregation rules. Therefore, before developing examples of how to use judgment aggregation in multiagent systems, we needed to develop judgment aggregation rules and a method for distinguishing among them. The results of this thesis are therefore not only in the filed of multiagent systems but also in the field of judgment aggregation theory. The two different types of groups we consider pose two different requirements for aggregation rules. Consensual groups need rules that produce decisions, which minimize the discrepancy with each individual opinion.

Hierarchical groups need rules that produce decisions that capitalize on the expertise of the group members. We therefore constructed two classes of rules: rules based on minimization and weighted distance-based rules.

In the first class of rules we constructed ten rules that generate decisions adherent to the majority: R_{MSA}, R_{MCSA}, R_{MWA}, R_{RA}, R_{Y}, R_{IY}, R_{RY}, R_{MR}, R_{MCR} and R_{MNAC}. We compare the decisions produced by one rule with the decisions produced by another, but in the same manner we compare our rules and existing rules in judgment aggregation. A summary of this analysis is given in Table 2.4.

In the second class of rules we start from the distance-based rules presented by (Pigozzi, 2006) and generalize them to rules that handle a richer structure of judgments, in particular various types of weights associated with the judgments. Apart from the sequential aggregation considered in (Li, 2010), no judgment aggregation rules have been proposed for aggregating ternary (or multi-valued) judgments. In (Dokow and Holzman, 2010b; Dietrich, 2007; Pauly and van Hees, 2006; Gärdenfors, 2006) frameworks for multi-valued judgment aggregation are considered, but no rules are proposed. Aggregation of weighted judgments has not been considered in the literature.

In belief merging weights associated with agents have already been considered in (Revesz, 1995) and merging multi-valued propositions has been considered in (Konieczny and Marquis, 2002; Condotta et al., 2008; Coste-Marquis et al., 2007). The novelty of our rules from the viewpoint of belief merging is in added possibility to assign weights to beliefs. Namely, the weight of a belief depends not only on the agent that holds the belief but also on which belief it is.

We analyze the complexity-theoretic properties of the weighted-distance based rules we propose. Much attention has been devoted to various complexity-theoretic aspects of voting rules, in particular to the problem of winner determination. The winner determination problem is the problem of determining if a given candidate is a winner for a given profile of votes when voting rule F is applied. In judgment aggregation the "winner" determination problem, *i.e.*, given a judgment and a profile of judgments determine if this judgment is among the selected collective judgments by judgment aggregation rule F, has only been consider in (Endriss et al., 2010b) for two judgment aggregation rules. The complexity-theoretic analysis of the winner determination problem is used as an indicator of the computational efficiency of a particular rule or aggregation operator. Certain complexity-theoretic aspects, corresponding to the winner determination problem as defined in (Endriss et al., 2010b), of belief merging operators have been considered in (Konieczny et al., 2004). Our complexity results are usable in a belief merging context as well.

8.1.2 Selecting judgment aggregation rules

To be able to distinguish among judgment aggregation rules within the same class, we need to consider which structural and relational properties are satisfied by these rules. Each author, or group of authors, that works in judgment aggregation theory has proposed their own framework for judgment aggregation, defining properties of judgment aggregation rules in it. As a consequence, we first needed to construct a general enough judgment aggregation framework common for both classes of rules and then construct the corresponding definitions of judgment aggregation rules within this framework. Only then were we able to analyze our rules with respect to the common rule properties considered in judgment aggregation theory.

Tables 4.14 and 4.15 summarize which of our rules satisfy which property. We repeat these tables here.

	Majority Preservation	Weak Unanimity	Strong Unanimity	IR-s	S-i
R_Y	✓	✓	✓	no	no
R_{MSA}	✓	✓	no	✓	no
R_{MCSA}	✓	no	no	✓	no
R_{MWA}	✓	no	no	✓	no
R_{RA}	✓	✓	✓	✓	no
$R^{d_H,max}$	no	no	no	no	no
R_{RY}	✓	✓	✓	no	no
R_{MNAC}	✓	no	no	no	no

Property	Satisfied	Not satisfied
Unanimity	(d,Σ) (d,max) $(d,Gmax)$ (d,Π^*)	
Weak unanimity principle	$(d_D,\odot)$	(d_i,Σ) (d_i,max) $(d_i,Gmax)$ (d_i,Π^*) $d_i \in \{d_H,d_T\}$
Strong unanimity principle	$(d_D,\odot)$	
Majoritarian	(d,Σ) (d,Π^*)	(d,max) $(d,Gmax)$
Majority-preserving	(d_H,Σ) (d_T,Σ)	(d_D,Σ) (d,max) $(d,Gmax)$ (d,Π^*)
IR	$(d,\odot)$	
S-s	(d,Σ) (d,max) $(d,Gmax)$ (d,Π^*)	

We analyzed our rules mainly for the judgment aggregation properties that already exist in the literature. These rules have been designed while searching for a minimal set of desirable properties that characterize a judgment aggregation rule, and consequently it is unsurprising that these are properties that are desirable in most contexts. We need the properties to distinguish among rules and to this end we need to develop and study more rule aggregator properties. We commence this line of research in judgment aggregation theory by defining five new desirable properties for rule aggregators, Section 4.10.

8.1.3 Applying judgment aggregation rules

As an example of hierarchical groups we consider a team of agents that solves problems in a changing environment. We use judgment aggregation to extend a model for reaching satisficing decisions developed in experimental psychology. The agent decision-maker considers one option at the time, choosing an option if it satisfies a combination of sufficiently relevant criteria, or cues as they are called in experimental psychology. The decision-maker aggregates the opinions of other agents to determine if these relevant criteria are satisfied, but also if the choosing of a particular option is supported. We propose our decision-making solution for incident management teams, however the same approach can be applied in various other contexts. One example is a recommendation system that uses the content of multiple web pages and other sources to recommend a product or a service to customers. Based on the criteria set by the customer, the system considers the sources that give information on the criteria, the reputation of the source and whether the source endorses the product/service in question.

As an example of consensual groups we consider a group of agents that needs to agree on what its intentions are. We propose a judgment aggregation based method for agreeing on which group goals to pursue. Our method is intended for groups that engage in joint activity when it is necessary that the group to present itself as a single whole from the point of view of beliefs and goals. The requirement that the group presents itself as a rational entity that has goals justified by the beliefs it holds, and is able to revise these goals under the light of new information, was held by (Tuomela and Miller, 1992) and adopted in agent theory by (Boella and van der Torre, 2007) and (Lorini and Longin, 2008). Our proposal to formalize group intentions as the goals on which the group agrees to pursue and is jointly committed to pursuing, can be applied, for example, in an open-source project, where several people have to discuss online to agree on which is their position on issues and which is their goal.

8.2 Other examples of using judgment aggregation

The thesis of (Ganesan, 2011) tests the use of judgment aggregation for the cooperative anchoring problem (LeBlanc and Saffiotti, 2008; Coradeschi and Loutfi, 2008) on NAO robots controlled by agents written in the GOAL programing language (Hindriks, 2010).

The anchoring, or symbol grounding, problem (Harnad, 1990) is the problem of assigning meaning to abstract symbols. This problem is considered as solved, namely that "we now understand enough to create systems in which groups of agents self-organizes a symbolic communication system that is grounded in their interactions with the world, and these systems may act as models to understand how humans manage to self-organize their communication systems." (Steels, 2008). Combining perceptual information may be used to ground a symbol. *E.g.,* the symbol p denoting "there is a chair in the room", can be grounded to true or false by considering vision, pattern recognition and sonar readings. Let as call symbols such as p, which can be grounded by using only percepts, a level- 0 symbol. A robot would also need to ground more complex symbols, such as q denoting "Room E112 is a meeting room", which would require fusing not only percepts but also level- 0 symbols.

Collective robotics is a field of research that is concerned with the development and use of robotic teams for performing various tasks. A team of robots needs to collectively ground and share symbols. There are two basic uses for social symbol grounding: the first is to establish grounding that will be used by the team when performing their tasks; the second to be used by a robot that grounds a symbol for individual use by considering the groundings made by several other robots.

The grounding of symbols higher than what we call here level- 0 is the problem of assigning values to a set of logically related issues. The social grounding of such symbols is the problem of aggregating the values assigned to the issues by different robots. This is a problem that can be modeled in judgment aggregation. Ganesan (2011) uses an example of a level- 0 symbol that needs to be socially grounded based on a set of percepts. She implements the premise-based and the conclusion-based procedure as described in (Kornhauser and Sager, 1993) and identifies the premise-based procedure as better at truth tracking. She also considers the distance-based procedure, as described in (Endriss et al., 2010b), for the obtained profiles. The premise-based procedure outperforms the distance-based procedure one in the case of level-0 symbols.

8.3 Future work

In computational contexts, as well in (human) society, there are many problems that require collectively binding decisions to be generated. Due to the variety of contexts in which these problems occur in society, social choice theory has been an active and not yet exhausted, research area for the last three centuries. An essential step towards advancing the use of judgment aggregation in multiagent systems is to look for properties that more finely distinguish among aggregation rules than the properties that are currently considered as desirable in judgment aggregation theory. Furthermore we need to search for, and characterize, collective decision problems in multiagent systems and study how their characteristics reflect into (un)desirable judgment aggregation properties.

In all the similarities between society and computational contexts, the requirements and constraints of the collective decision-making in these two contexts are different. How social choice rules are used in society and multiagent systems is different: while people need to rely on them sporadically and only when consensus fails to emerge, artificial agents need to use them for every single group decisions they need to make.

It is often insufficient to analyze the computational properties of existing social choice rules, but one needs to design more. This is the first thing we observed when starting to explore the possible use of judgment aggregation in multiagent systems. For instance, the premise-based and conclusion-based rules are enough of rule options in collegiate courts. However, a recommendation service needs to consider a richer structure of judgments and agendas cannot necessarily be partitioned into premises and conclusions. The collection of judgment aggregation rules that we proposed in this thesis is not exhaustive.

The differences between social and computational contexts also apply a different approach to a complexity-theoretic analysis of the aggregation rules. The problem of confirming that a judgment or a judgment set is selected collective, by a given rule for a given profile, is relevant in a human society. Namely, after a consensus starts to emerge, a group casts judgments and only needs to verify that this consensus is the group decision, with respect to the individual judgments and the rule used. Artificial agents do not start a group decision-reaching process by an informal chat, *i.e.*, either argumentation is not combined with voting as its often the case in naturalistic settings. In multiagent systems contexts it is more relevant to consider the complexity of finding a collective judgment set rather then confirming that a given judgment set is the collective one.

The rules designed and represented here are irresolute by construction. Designing rules in this manner is necessary if one wishes to avoid impossibility and not impose domain restrictions. However, a group usually needs only one decision, *i.e.*, judgment set. Creating resolute judgment aggregation procedures can be done by implementation rather than by construction. How a social choice rule is implemented is an important question in formal voting contexts (Dasgupta et al., 1979; Repullo, 1985; Maskin, 1999; Palfrey, 2002; Serrano, 2004). Implementation of judgment aggregation rules is an important issue in multiagent systems contexts. The choice of implementation unveils the amount and nature of resources needed for judgment aggregation, but also allows for certain behavior of the agents to be enforced or discouraged. Answering the question of implementation will also shed light on the relations between game theory and judgment aggregation, which are unexplored but bound to exist. As mentioned, committees in society usually discuss before voting. For multiagent systems this discussion segment could be modeled as an argumentation based dialog game and combined with the aggregation rules proposed in this thesis.

We considered hierarchical groups verses consensual groups. A natural option is embodied agents verses software agents. The thesis of (Ganesan, 2011) tests judgment aggregation on a social symbol grounding problem of a level- 0 symbol for the premise-based and conclusion-based procedure, but the same experiment can be ran on the remainder of the aggregation rules. To obtain meaningful results one needs to consider a larger number of robots and a bigger agenda. The technical challenges made it difficult to extend the experiment into social symbol grounding for symbols above level- 0. These challenges were introduced by the robot-agent and robot-robot communication that needed to be technically solved. Since these issues are resolved, a natural continuation of this project is to design examples and test the performance of judgment aggregation for symbols above level- 0. In general, social robotics is an emerging area of research in which one can expect many group decision-reaching problems to emerge.

Bibliography

Alchourrón, C. E., Gärdenfors, P., and Makinson, D. (1985). On the logic of theory change: Partial meet contraction and revision functions. *The Journal of Symbolic Logic*, 50(2):510–530. → pages 178

Arrow, K. (1963). *Social choice and individual values*. Cowles Foundation Monograph Series. → pages 7, 85, 86

Arrow, K. (1986). Rationality of self and others in an economic system. *Journal of Business*, 59(4):S385–99. → pages 142

Arrow, K., Sen, A., and Suzumura, K. (2002). *Handbook of Social Choice and Welfare*, volume 1. Elsevier. → pages 2

Bartholdi, J., Tovey, C. A., and Trick, M. A. (1989). Voting schemes for which it can be difficult to tell who won the election. *Social Choice and Welfare*, 6:157–165. → pages 3

Beer, M., D'inverno, M., Luck, M., Jennings, N., Preist, C., and Schroeder, M. (1999). Negotiation in multi-agent systems. *Knowledge Engineering Review*, 14:285–289. → pages 4

Benamara, F., Kaci, S., and Pigozzi, G. (2010). Individual opinions-based judgment aggregation procedures. In Torra, V., Narukawa, Y., and Daumas, M., editors, *Modeling Decisions for Artificial Intelligence*, volume 6408 of *Lecture Notes in Computer Science*, pages 55–66. Springer Berlin / Heidelberg. → pages 180, 181

Bienvenu, M., Lang, J., and Wilson, N. (2010). From preference logics to preference languages, and back. In Lin, F., Ulrike Sattler, U., and Truszczynski, M., editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the 12th International Conference*. → pages 3

Blyth, C. R. (1972). On Simpson's paradox and the sure-thing principle. *Journal of the American Statistical Association*, 67(338):364–366. → pages 120

Bochvar, D. (1938). Ob odnom tréhznačhom isčislénii i égo priménéii k analizu paradoksov kalssičéskogo rassirénnogo funkcional'nogo isčislenija (tr. on a 3-valued logical calculus and its application to the analysis of contradictions)). *Maténatičeskij sbornik*, 4:287–308. → pages 147, 148

Boella, G., da Costa Pereira, C., Pigozzi, G., Tettamanzi, A., and van der Torre, L. (2007). Choosing your beliefs. In *Normative Multi-agent Systems*. http://drops.dagstuhl.de/opus/volltexte/2007/938/. → pages 159

Boella, G., Pigozzi, G., Slavkovik, M., and van der Torre, L. (2011a). Group intention is social choice with commitment. In De Vos, M., Fornara, N., Pitt, J., and Vouros, G., editors, *Coordination, Organizations, Institutions, and Norms in Agent Systems VI*, volume 6541 of *Lecture Notes in Computer Science*, pages 152–171. Springer Berlin / Heidelberg. → pages 18

Boella, G., Pigozzi, G., Slavkovik, M., and van der Torre, L. (2011b). A satisficing agreements model. In van Riemsdijk, M. B. and Salceda, J., editors, *Proceedings of the 13th International Workshop on Coordination, Organization, Institutions and Norms (COIN@WI-IAT)*, page online proceedings. → pages 17, 18

Boella, G. and van der Torre, L. (2007). The ontological properties of social roles in multi-agent systems: Definitional dependence, powers and roles playing roles. *Artificial Intelligence and Law Journal (AILaw)*, 15(3):201–221. → pages 188

Borkowski, L., editor (1970). *Selected Works*. Noth-Holland and Warsaw. → pages 147

Bouveret, S. and Lang, J. (2005). Efficiency and envy-freeness in fair division of indivisible goods: logical representation and complexity. In *International Joint Conference on Artificial Intelligence*, pages 935–940. → pages 3

Brams, S. and Fishburn., P. (2004). Chapter 4: Voting procedures. In Arrow, K., Sen, A., and Suzumura, K., editors, *Handbook of Social Choice and Welfare*. Elsevier. → pages 16, 22

Brams, S., Kilgour, D., and Sanver, R. (2007). A minimax procedure for electing committees. *Public Choice*, 132(3-4):401–420. → pages 35

Brams, S. J. and Taylor, A. (1996). *Fair Division: From Cake-Cutting to Dispute Resolution*. Cambridge University Press. → pages 50

Bratman, M. E. (1993). Shared intention. *Ethics*, 104(1):97–113. → pages 162

Brown, C. B. (1990). Optimizing and satisficing. *Structural Safety*, 7(2-4):155 – 163. → pages 142

Burke, C. S., Stagl, K. C., Salas, E., and Pierce, L.; Kendall, D. (2006). Understanding team adaptation: A conceptual analysis and model. *Journal of Applied Psychology*, 91(6):1189–1207. → pages 143, 154

Cadoli, M. and Schaerf, M. (1996). On the complexity of entailment in propositional multivalued logics. *Annals of Mathematics and Artificial Intelligence*, 18:29–50. → pages 147

Cariani, F., Pauly, M., and Snyder, J. (2008). Decision framing in judgment aggregation. *Synthese*, 163(1):1–24. → pages 123

Castelfranchi, C. and Paglieri, F. (2007). The role of beliefs in goal dynamics: Prolegomena to a constructive theory of intentions. *Synthese*, 155:237–263. → pages 159, 163

Chapman, B. (1998). More easily done than said: Rules, reasons and rational social choice. *Oxford Journal of Legal Studies*, 18(2):pp. 293–329. → pages 16

Chapman, B. (2002). Rational aggregation. *Politics, Philosophy and Economics*, 1(3):337–354. → pages 91

Chellas, B. F. (1980). *Modal Logic: An Introduction*. Cambridge University Press, Cambridge. → pages 163

Chevaleyre, Y., Dunne, P., Endriss, U., Lang, J., Maudet, N., and Rodríguez-Aguilar, J. (2005). Multiagent resource allocation. *Knowledge Engineering Review. Special Issue on the AgentLink III Technical Forums*, 20(2):143–149. → pages 3

Chevaleyre, Y., Endriss, U., Lang, J., and Maudet, N. (2007). A short introduction to computational social choice. In *Proceedings of the 33rd conference on Current Trends in Theory and Practice of Computer Science*, SOFSEM '07, pages 51–69, Berlin, Heidelberg. Springer-Verlag. → pages 3, 74, 137

Cohen, P. R. and Levesque, H. (1990). Intention is choice with commitment. *Artificial Intelligence*, 42(2-3):213–261. → pages 14, 15, 159, 161

Condorcet, M. M. d. (1785). *Essai sur l'Application de l'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix*. Paris. → pages 95

Condotta, J., Kaci, S., and Schwind, N. (2008). A framework for merging qualitative constraints networks. In *The International Conference of the Florida Artificial Intelligence Research Society*, FLAIRS, pages 586–591. → pages 12, 179, 186

Conitzer, V. and Sandholm, T. (2002a). Complexity of manipulating elections with few candidates. In *Proceedings of the 18th National Conference on Artificial Intelligence and 18th Conference on Innovative Applications of Artificial Intelligence*, AAAI/IAAI, pages 314–319. American Association of Artificial Intelligence. → pages 3

Conitzer, V. and Sandholm, T. (2002b). Vote elicitation: Complexity and strategy-proofness. In *Proceedings of the 18th National Conference on Artificial Intelligence and Fourteenth Conference on Innovative Applications of Artificial Intelligence*, AAAI/IAAI, pages 392–397. American Association of Artificial Intelligence. → pages 3, 175

Conitzer, V. and Sandholm, T. (2005). Communication complexity of common voting rules. In *ACM Conference on Electronic Commerce*, pages 78–87. → pages 175

Coradeschi, S. and Loutfi, A. (2008). A review of past and future trends in perceptual anchoring. In Fritzsche, P., editor, *Tools in Artificial Intelligence*, chapter 15. InTech. → pages 188

Coste-Marquis, S., Devred, C., and Konieczny, S. (2007). On the merging of Dung's argumentation systems. *Artificial Intelligence*, 171(10-15):730–753. → pages 179, 186

Dasgupta, P., Hammond, P., and Maskin, E. (1979). The implementation of social choice rules: Some general results on incentive compatibility. *The Review of Economic Studies*, 46(2):pp. 185–216. → pages 190

Deza, M. and Deza, E. (2009). *Encyclopedia of Distances*. Springer. → pages 58

Dietrich, F. (2006a). Aggregation theory and the relevance of some issues to others. Technical report, Maastricht research school of Economics of Technology and Organizations. → pages 86

Dietrich, F. (2006b). Judgment aggregation: (im)possibility theorems. *Journal of Economic Theory*, 126:286–298. → pages 123, 181

Dietrich, F. (2007). A generalised model of judgment aggregation. *Social Choice and Welfare*, 28(4):529–565. → pages 16, 52, 53, 71, 179, 180, 186

Dietrich, F. (2010). The possibility of judgment aggregation on agendas with subjunctive implications. *Journal of Economic Theory*, 145(2):603 – 638. → pages 22, 177

Dietrich, F. and List, C. (2005). Strategy-proof judgment aggregation. STICERD - Political Economy and Public Policy Paper Series 09, Suntory and Toyota International Centres for Economics and Related Disciplines, LSE. → pages 86, 87, 91, 112, 181

Dietrich, F. and List, C. (2007a). Arrow's theorem in judgment aggregation. *Social Choice and Welfare*, 29(1):19–33. → pages 95, 166

Dietrich, F. and List, C. (2007b). Judgment aggregation by quota rules: Majority voting generalized. *Journal of Theoretical Politics*, 4(19):391 – 424. → pages 16, 17, 22, 74, 100, 177

Dietrich, F. and List, C. (2008a). Judgment aggregation without full rationality. *Social Choice and Welfare*, 31(1):15–39. → pages 85, 86, 112, 181

Dietrich, F. and List, C. (2008b). A liberal paradox for judgment aggregaton. *Social Choice and Welfare*, 31(1):59–78. → pages 86, 107, 108, 181

Dietrich, F. and List, C. (2010). Majority voting on restricted domains. *Journal of Economic Theory*, 145(2):512–543. → pages 91

Dietrich, F. and Mongin, P. (2010). The premisse-based approach to judgment aggregation. *Journal of Economic Theory*, 145(2):562–582. → pages 16, 17, 21, 62, 72, 177

Dijkstra, E. W. (1959). A note on two problems in connection with graphs. *Numerische Mathematik*, 1:269–271. → pages 76

Dodgson, C. (1876). *A Method for Taking Votes on More than Two Issues*. Clarendon Press. → pages 31

Dokow, E. and Holzman, R. (2010a). Aggregation of binary evaluations. *Journal of Economic Theory*, 145(2):495 – 511. → pages 86

Dokow, E. and Holzman, R. (2010b). Aggregation of binary evaluations with abstentions. *Journal of Economic Theory*, 145(2):544 – 561. → pages 52, 89, 179, 186

Duddy, C. and Piggins, A. (2011). A measure of distance between judgment sets. (working paper). → pages 61, 76

Dunin-Keplicz, B. and Verbrugge, R. (2010). *Teamwork in Multi-Agent Systems: A Formal Approach*. Wiley and Sons. → pages 5, 18, 160, 162, 182

Elkind, E., Faliszewski, P., and Slinko, A. (2009). On distance rationalizability of some voting rules. In *Proceedings of the 12th Conference on Theoretical Aspects of Rationality and Knowledge (TARK '09)*, pages 108–117. → pages 31

Endriss, U., Grandi, U., and Porello, D. (2010a). Complexity of judgment aggregation: Safety of the agenda. In *Proceedings of the 9th International Joint Conference on Autonomous Agents and Multiagent Systems*, AAMAS-2010, pages 359–366. IFAAMAS. Also presented at SCW-2010. → pages 16, 17, 77, 78, 80, 81

Endriss, U., Grandi, U., and Porello, D. (2010b). Complexity of winner determination and strategic manipulation in judgment aggregation. In *Proceedings of the 3rd International Workshop on Computational Social Choice (COMSOC-2010)*. University of Düsseldorf. → pages 25, 30, 34, 35, 56, 72, 74, 78, 87, 178, 186, 189

Everaere, P., Konieczny, S., and Marquis, P. (2007). The strategy-proofness landscape of merging. *Journal of Artificial Intelligence Research (JAIR)*, 28:49–105. → pages 87

Fishburn, P. C. and Brams, S. J. (1983). Paradoxes of preferential voting. *Mathematics Magazine*, 56(4):207–214. → pages 125

Ganesan, V. (2011). Using judgement aggregation for social symbol grounding on the NAO robots. Master's thesis, University of Luxembourg. → pages 88, 188, 189, 190

Gärdenfors, P. (2006). A representation theorem for voting with logical consequences. *Economics and Philosophy*, 22(2):181–190. → pages 52, 179, 186

Gilbaud, G. (1966). Theories of the general interest, and the logical problem of aggregation. In Lazarsfeld, P. and Henry, N., editors, *Readings in Mathematical Social Science*, pages 262–307. MIT Press, Cambridge, MA, USA. → pages 6, 53

Gilbert, M. P. (1987). Modeling Collective Belief. *Synthese*, 73:185–204. → pages 18, 160, 163, 165

Gilbert, M. P. (2002). Belief and acceptance as features of groups. *Protosociology: An International Journal of Interdisciplinary Research*, 16:35–69. → pages 18, 160, 163

Gilbert, M. P. (2007). Acting together, joint commitment, and obligation. *Philosophische Analyse/Philosophical Analysis*. → pages 18, 160, 170

Gilbert, M. P. (2009). Shared Intention and Personal Intentions. *Philosophical Studies, 2009*. → pages 5, 15, 18, 160, 162

Goldman, A. I. (1987). Foundations of social epistemics. *Synthese*, 73(1):pp. 109–144. → pages 15

Grabisch, M., Marichal, J.-L., Mesiar, R., and Pap, E. (2009). *Aggregation Functions*. Cambridge University Press, 1st edition. → pages 57

Grandi, U. and Endriss, U. (2010). Lifting rationality assumptions in binary aggregation. In Fox, M. and Poole, D., editors, *Proceedings of the 24th AAAI Conference on Artificial Intelligence*, AAAI'10. AAAI Press. → pages 85, 177, 181

Grandi, U. and Endriss, U. (2011). Binary aggregation with integrity constraints. In Walsh, T., editor, *Proceedings of the 22nd International Joint Conference on Artificial Intelligence*, IJCAI'11, pages 204–209. → pages 85, 86, 89, 181

Greiner, R., Hayward, R., Jankowska, M., and Molloy, M. (2006). Finding optimal satisficing strategies for and-or trees. *Artificial Intelligence*, 170(1):19 – 58. → pages 142

Grossi, D., Pigozzi, G., and Slavkovik, M. (2009). White manipulation in judgment aggregation. In *Proceedings of the 21st Benelux Conference on Artificial Intelligence*, page online, Eindhoven, The Netherlands. → pages 18

Grosz, B. and Hunsberger, L. (2007). The dynamics of intention in collaborative activity. *Cognitive Systems Research*, 7(2-3):259–272. → pages 160, 182, 183

Güth, W. (2010). Satisficing and (un)bounded rationality—a formal definition and its experimental validity. *Journal of Economic Behavior and organization*, 73(3):308 – 316. → pages 142

Hakli, R. (2006). Group beliefs and the distinction between belief and acceptance. *Cognitive Systems Research*, 7(2-3):286 – 297. Cognition, Joint Action and Collective Intentionality. → pages 159, 163

Haller, H. (1985). The principal-agent problem with a satisficing agent. *Journal of Economic Behavior and Organization*, 6(4):359 – 379. → pages 142

Hardy-Vallèe, B., editor (2007). *Cognitive Decision-Making: Empirical and Foundational Issues*, Cambridge. Cambridge Scholars Press. → pages 133

Harnad, S. (1990). The symbol grounding problem. *Physica D*, 42:335–346. → pages 188

Hartmann, S., Pigozzi, G., and Sprenger, J. (2010). Reliable methods of judgment aggregation. *Journal of Logic and Computation*, 20(2):603–617. → pages 88

Hemaspaandra, E., Hemaspaandra, L. A., and Rothe, J. (1997). Exact analysis of dodgson elections: Lewis carroll's 1876 voting system is complete for parallel access to NP. *Journal of the Association for Computing Machinery*, 44(6):806–825. → pages 3

Hindriks, K. (2010). *GOAL Programming Guide (Draft)*. Delft University of Technology. → pages 188

Hollenbeck, J. R., Ilgen, D. R., Sego, D. J., Hedlund, J., Major, D. A., and Phillips, J. (1995). Multilevel theory of team decision making: Decision performance in teams incorporating distributed expertise. *Journal of Applied Psychology*, 80(2):292 – 316. → pages 134

Hulin, C. and Ilgen, D. R. (2000). Introduction to computational modeling in organizations: the good that modeling does. In Ilgen, D. R. and Hulin, C., editors, *Computational Modeling of Behavior in Organizations: The Third Scientific Discipline*, pages 3–18. American Psychological Association. → pages 134

Humphrey, S., Hollenbeck, J. R., Meyer, C., and Ilgen, D. R. (2002). Hierarchical team decision-making. In Ferris, G. and Martocchio, J., editors, *Research in Personnel and Human Resources Management*, pages 175–214. Elsevier. → pages 134

Hunsberger, L. (2002). *Group Decision Making and Temporal Reasoning*. PhD thesis, Harvard University, Available as Harvard Technical Report TR-05-02. → pages 183

Icard, T., Pacuit, E., and Shoham, Y. (2010). Joint revision of beliefs and intention. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Twelfth International Conference*, KR'10, http://aaai.org/ocs/index.php/KR/KR2010/paper/view/1297. AAAI Press. → pages 183

Ilgen, D. R., Hollenbeck, J. R., Johnson, M., and Jundt, D. (2005). Teams in organizations: From input-process-output models to IMOI models. *Annual Review of Psychology*, 56(1):517–543. → pages 134

J., L. and L., M. (1990). Satisficing in knowledge-based systems. *Data and Knowledge Engineering*, 5(2):139 – 166. → pages 142

Jennings, N. R. (1995). Controlling cooperative problem solving in industrial multi-agent systems using joint intentions. *Artificial Intelligence Journal*, 75(2):195–240. → pages 160, 162, 182

Kaufman, B. E. (1990). A new theory of satisficing. *Journal of Behavioral Economics*, 19(1):35 – 51. → pages 142

Kleene, S. C. (1938). On notation for ordinal numbers. *The Journal of Symbolic Logic*, 3(4):150–155. → pages 64, 147, 148, 179

Klein, G. (1999). *Sources of Power: How People Make Decisions*. The MIT Press. → pages vii, 134, 137, 144, 156

Klein, G., Calderwood, R., and Clinton-Cirocco, A. (2010). Rapid decision making on the fire ground: The original study plus a postscript. *Journal of Cognitive Engineering and Decision Making*, 4:186–209. → pages vii, 13, 134, 138

Konieczny, S., Lang, J., and Marquis, P. (2004). DA^2 merging operators. *Artificial Intelligence Journal*, 157:45–79. → pages 12, 18, 52, 57, 58, 178, 186

Konieczny, S. and Marquis, P. (2002). Three-valued logics for inconsistency handling. In Flesca, S., Greco, S., Ianni, G., and Leone, N., editors, *Logics in Artificial Intelligence*, volume 2424 of *Lecture Notes in Computer Science*, pages 332–344. Springer Berlin / Heidelberg. → pages 180, 186

Konieczny, S. and Pino-Pérez, R. (1999). Merging with integrity constraints. *Lecture Notes in Computer Science*, 1638/1999:233–244. → pages 12, 52, 56, 57, 58, 80, 94, 128

Konieczny, S. and Pino-Pérez, R. (2002). Merging information under constraints: a logical framework. *Journal of Logic and Computation*, 12(5):773–808. → pages 52, 94, 99

Konieczny, S. and Pino-Pérez, R. (2011). Logic based merging. *Journal of Philosophical Logic*, 40(2):239–270. → pages 35

Konieczny, S. and Pino-Pérez, R. R. (2005). Propositional belief base merging or how to merge beliefs/goals coming from several sources and some links with social choice theory. *European Journal of Operational Research*, 160(3):785–802. → pages 94

Kornhauser, L. and Sager, L. (1986). Unpacking the court. *Yale Law Journal*, 96:82–117. → pages 6

Kornhauser, L. and Sager, L. (1993). The one and the many: Adjudication in collegial courts. *California Law Review*, 81:1–51. → pages 16, 21, 71, 177, 189

Lang, J., Pigozzi, G., Slavkovik, M., and van der Torre, L. (2011). Judgment aggregation rules based on minimization. In *Theoretical Aspects of Rationality and Knowledge*, TARK'11, pages 238–246. → pages 17, 18

LeBlanc, K. and Saffiotti, A. (2008). Cooperative anchoring in heterogeneous multi-robot systems. In Mataric, M., Schenker, P., Schaal, S., and Sukhatme, G., editors, *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*, pages 3308 –3314. IEEE Xplore. → pages 188

Levesque, H. J., Cohen, P. R., and Nunes, J. H. T. (1990). On acting together. In *Proceedings of the 8th National Conference on Artificial Intelligence*, AAAI'90, pages 94–99. → pages 162

Li, N. (2010). Decision paths in sequential non-binary judgment aggregation. Technical report, Universitat Autònoma de Barcelona. → pages 22, 52, 74, 90, 177, 181, 186

List, C. (2004a). A model of path-dependence in decisions over multiple propositions. *The American Political Science Review*, 98(3):495–513. → pages 22, 74, 177

List, C. (2004b). A model of path-dependence in decisions over multiple propositions. *American Political Science Review*, 3(98):495 – 513. → pages 30

List, C. (2005). Group knowledge and group rationality: a judgment aggregation perspective. *Episteme*, 2(1):pp. 25–38. → pages 15

List, C. and Pettit, P. (2002). Aggregating sets of judgments: An impossibility result. *Economics and Philosophy*, 18(01):89–110. → pages 6, 16, 53, 86, 90, 91

List, C. and Puppe, C. (2004). Aggregating sets of judgments: Two impossibility results compared. *Synthese*, 104(1-2):207–235. → pages 6

List, C. and Puppe, C. (2009). Judgment aggregation: A survey. In Anand, P., Puppe, C., and Pattanaik, P., editors, *Oxford Handbook of Rational and Social Choice*. Oxford. → pages 2, 7, 86, 90, 91, 107, 112, 137, 181

List, L. and Polak, B. (2010). Introduction to judgment aggregation. *Journal of Economic Theory*, 145(2):441 – 466. → pages 6, 16, 86, 89, 91, 127, 128, 177, 178

Lorini, E. and Longin, D. (2008). A logical account of institutions: From acceptances to norms via legislators. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Eleventh International Conference*, KR'08, pages 38–48. → pages 188

Lorini, E., Longin, D., Gaudou, B., and Herzig, A. (2009). The logic of acceptance. *Journal of Logic and Computation*, 19(6):901–940. → pages 163, 165

Łukasiewicz, J. (1920). O logice trójwartościowej. *Ruch Filozoficzny*, 5:170–171. → pages 89, 147, 148, 179

Maskin, E. (1999). Nash equilibrium and welfare optimality. *Review of Economic Studies*, 66(1):23–38. → pages 190

Matsuda, T. and Takatsu, S. (1979). Characterization of satisficing decision criterion. *Information Sciences*, 17(2):131 – 151. → pages 142

Maudet, N. (2010). Reaching agreement in multiagent systems. Habilitation, Université Paris Dauphine. → pages 3

Meijers, A. (2002). Collective agents and cognitive agents. *Protosociology. Special Issue Understanding the Social:New Perspectives from Epistemology*, 16:70–86. → pages 15, 160, 163

Miller, M. (2008). Judgment aggregation and subjective decision-making. *Economics and Philosophy*, 24(02):205–231. → pages 16, 64, 180

Miller, M. and Osherson, D. (2009). Methods for distance-based judgment aggregation. *Social Choice and Welfare*, 32(4):575–601. → pages 16, 22, 25, 27, 28, 34, 35, 50, 177

Monahan, G. E. (1982). A survey of partially observable markov decision processes: Theory, models, and algorithms. *Management Science*, 28(1):pp. 1–16. → pages 15

Mongin, P. (2008). Factoring out the impossibility of logical aggregation. *Journal of Economic Theory*, 141(1):100–113. → pages 21, 72

Moulin, H. (1988). Condorcet's principle implies the no show paradox. *Journal of Economic Theory*, 45(1):53–64. → pages 126

Moulin, H. (1991). *Axioms of cooperative decision making*. Number 15 in Econometric Society Monographs. Cambridge University Press. → pages 121

Nash, J. R. (2003). A context-sensitive voting protocol paradigm for multimember courts. *Stanford Law Review*, 56:75–159. → pages 12

Nehring, K. and Pivato, M. (2011). Majority rule in the absence of a majority. Presented at the New Developments in Judgement Aggregation and Voting Theory Workshop. → pages 50, 52, 85, 178, 181

Nehring, K., Pivato, M., and Puppe, C. (2011). Condorcet admissibility: Indeterminacy and path-dependence under majority voting on interconnected decisions. http://mpra.ub.uni-muenchen.de/32434/. → pages 50, 52, 85, 96, 178, 181

Nehring, K. and Puppe, C. (2005). The structure of strategy-proof social choice. part II: Non-dictatorship, anonymity and neutrality. *Unpublished manuscript, University of Karlsruhe*. → pages 91

Nehring, K. and Puppe, C. (2010a). Abstract Arrowian aggregation. *Journal of Economic Theory*, 145(2):467–494. → pages 86

Nehring, K. and Puppe, C. (2010b). Justifiable group choice. *Journal of Economic Theory*, 145(2):583 – 602. → pages 85

Nurmi, H. (2004). Monotonicity and its cognates in the theory of choice. *Public Choice*, 121(1/2):pp. 25–49. → pages 112, 126

Nurmi, H. (2010). Voting theory. In Rios Insua, D. and French, S., editors, *e-Democracy*, volume 5 of *Advances in Group Decision and Negotiation*, pages 101–123. Springer Netherlands. → pages 2, 137

Ossowski, S. (2008). Coordination and agreement in multi-agent systems. In Klusch, M., Pechoucek, M., and Polleres, A., editors, *Cooperative Information Agents XII*, volume 5180 of *Lecture Notes in Computer Science*, pages 16–23. Springer Berlin Heidelberg. → pages 1, 4

Othman, A. and Sandholm, T. (2010). When do markets with simple agents fail? In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1 - Volume 1*, AAMAS '10, pages 865–872, Richland, SC. International Foundation for Autonomous Agents and Multiagent Systems. → pages 5

Palfrey, T. R. (2002). Implementation theory (chapter 61). In Aumann, R. and Hart, S., editors, *Handbook of Game Theory with Economic Applications*, volume 3, chapter 61, pages 2271 – 2326. Elsevier. → pages 190

Papadimitriou, C. H. (1994). *Computational Complexity*. Addison-Wesley. → pages 75, 80

Pauly, M. and van Hees, M. (2006). Logical constraints on judgement aggregation. *Journal of Philosophical Logic*, 35(6):569 – 585. → pages 16, 52, 53, 71, 90, 179, 186

Peterson, M. (2009). *An Introduction to Decision Theory*. Cambridge University Press, 1st edition. → pages 133, 141

Pigozzi, G. (2006). Belief merging and the discursive dilemma: an argument-based account to paradoxes of judgment aggregation. *Synthese*, 152(2):285–298. → pages 11, 12, 16, 22, 34, 127, 177, 178, 179, 186

Pigozzi, G., Slavkovik, M., and van der Torre, L. (2008a). Desirable tie-breaking rules in collective decision making. In Roy, B. and Paschos, V., editors, *Annales du Lamsade, Proceedings of the DIMACS-LAMSADE Workshop on Algorithmic Decision Theory*, pages 171–174, Paris, France. → pages 18

Pigozzi, G., Slavkovik, M., and van der Torre, L. (2008b). Independence in judgment aggregation. In *Proceedings of the 9th Ninth International Meeting of the Society for Social Choice and Welfare*, Montreal, Canada. → pages 18

Pigozzi, G., Slavkovik, M., and van der Torre, L. (2009). A complete conclusion-based procedure for judgment aggregation. In Rossi, F. and Tsoukiàs, A., editors, *Algorithmic Decision Theory, 1rst International Conference, ADT 2009, Venice, Italy, October 20-23, 2009. Proceedings*, volume 5783 of *Lecture Notes in Computer Science*, pages 1–13. Springer. → pages 17, 72, 109, 177

Pnueli, A. (1977). The temporal logic of programs. In *SFCS '77: Proceedings of the 18th Annual Symposium on Foundations of Computer Science*, pages 46–57, Washington, DC, USA. IEEE Computer Society. → pages 163

Post, E. L. (1921). Introduction to a general theory of elementary propositions. *American Journal of mathematics*, 43:163–185. → pages 64, 147, 179

Quinton, A. (1975). The presidential address: Social objects. *Proceedings of the Aristotelian Society*, 76:pp. 1–27+viii. → pages 159

Radner, R. (1975). Satisficing. *Journal of Mathematical Economics*, 2(2):253 – 262. → pages 142

Rahwan, I., Ramchurn, S. D., Jennings, N. R., Mcburney, P., Parsons, S., and Sonenberg, L. (2003). Argumentation-based negotiation. *Knowledge Engineering Review*, 18(3):343–375. → pages 137

Rao, A. S. and Georgeff, M. P. (1993). Intentions and rational commitment. In *In Proceedings of the First Pacific Rim Conference on Artificial Intelligence (PRICAI-90)*, pages 94–99. → pages 171, 174, 175

Rao, A. S., Georgeff, M. P., and Sonenberg, E. A. (1992). Social plans: a preliminary report. *Association for Computing Machinery Spcial Interest Group for Office Information Systems Bulletin*, 13:10–. → pages 162

Repullo, R. (1985). Implementation in dominant strategies under complete and incomplete information. *The Review of Economic Studies*, 52(2):223–229. → pages 190

Revesz, P. Z. (1995). On the semantics of arbitration. *International Journal of Algebra and Computation*, 7:133–160. → pages 12, 52, 179, 186

Roughgarden, T. (2010). Algorithmic game theory. *Commununications of the ACM*, 53(7):78–86. → pages 3

Roy, O. (2009a). A dynamic-epistemic hybrid logic for intentions and information changes in strategic games. *Synthese*, 171:291–320. → pages 183

Roy, O. (2009b). Intentions and interactive transformations of decision problems. *Synthese*, 169:335–349. → pages 183

Rubinstein, A. and Fishburn, P. C. (1986). Algebraic aggregation theory. *Journal of Economic Theory*, 38(1):63–77. → pages 6, 16, 53

Russell, S. and Norvig, P. (2010). *Artificial Intelligence: a modern approach*. Prentice Hall, 3rd edition edition. → pages 1, 3

Schild, K. (2000). On the relationship between BDI logics and standard logics of concurrency. *Autonomous Agents and Multi-Agent Systems*, 3(3):259–283. → pages 161, 163, 164

Schulze, M. (2003). A new monotonic and clone-independent single-winner election method. *Voting Matters*, 17:9–19. → pages 30

Serrano, R. (2004). The theory of implementation of social choice rules. *Society for Industrial and Applied Mathematics (SIAM) Review*, 46(3):377–414. → pages 190

Shinzo and Takatsu (1980). Decomposition of satisficing decision problems. *Information Sciences*, 22(2):139 – 148. → pages 142

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1):99–118. → pages 133, 141, 182

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, 63(2):129–138. → pages 15, 133, 141

Singh, M. P. (1990). Group intentions. In *Proceedings of the 10th International Workshop on Distributed Artificial Intelligence*, DAI'90. http://www.csc.ncsu.edu/faculty/mpsingh/papers/mas/dai-90.ps.gz. → pages 160, 182

Slavkovik, M. and Jamroga, W. (2011). Distance-based judgment aggregation of three-valued judgments with weights. In Elkind, E., Endriss, U., and Lang, J., editors, *Proceedings of the IJCAI Workshop on Social Choice and Artificial Intelligence*, pages 81–87. → pages 17, 18

Smith, J. (1973). Aggregation of preferences with variable electorate. *Econometrica*, 41(6):1027–1041. → pages 87, 120

Steels, L. (2008). The Symbol Grounding Problem Has Been Solved. So What's Next? In de Vega, M., editor, *Symbols and Embodiment: Debates on Meaning and Cognition*, chapter 12. Oxford University Press, Oxford. → pages 188

Stirling, W. and Nokleby, M. (2009). Satisficing coordination and social welfare for robotic societies. *International Journal of Social Robotics*, 1:53–69. → pages 133, 142, 182

Stirling, W. C. and Goodrich, M. A. (1999). Satisficing games. *Information Sciences*, 114(1-4):255 – 280. → pages 182

Tang, P. and Lin, F. (2009). Computer-aided proofs of Arrow's and other impossibility theorems. *Artificial Intelligence*, 173(11):1041 – 1053. → pages 3

Tideman, T. N. (1987). Independence of clones as a criterion for voting rules. *Social Choice and Welfare*, 4:185–206. → pages 87, 123

Tuomela, R. and Miller, K. (1992). Groups beliefs. *Synthese*, 91:285–318. → pages 188

Uckelman, J. and Endriss, U. (2010). Compactly representing utility functions using weighted goals and the max aggregator. *Artificial Intelligence Journal*, 174:1222–1246. → pages 183

Urquhart, A. (2001). Basic many-valued logic. In Gabbay, D. and Guenthener, F., editors, *Handbook of Philosophical Logic (Second edition)*, volume 2, pages 249–295. Kluwer Academic Publishers. → pages 54, 89, 146, 147

van der Hoek, W., Jamroga, W., and Wooldridge, M. (2007). Towards a theory of intention revision. *Synthese*, 155(2):265–290. → pages 182, 183

Walsh, T. (2008). Complexity of terminating preference elicitation. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems - Volume 2*, AAMAS '08, pages 967–974, Richland, SC. International Foundation for Autonomous Agents and Multiagent Systems. → pages 3

Wierzbicki, A. (1982). A mathematical basis for satisficing decision making. *Mathematical Modelling*, 3(5):391 – 405. → pages 142

Williams, H. (2009). *Logic and Integer programming*, volume 130 of *International Series in operational research and management science*. Springer. → pages 78

Wilson, R. (1975). On the theory of aggregation. *Journal of Economic Theory*, 10(1):89 – 99. → pages 6, 16, 53

Wolter, F. (1998). Fusions of modal logics revisited. In Kracht, M., de Rijke, M., and Zakharyaschev, M., editors, *Advances in Modal Logic' 96*, pages 361–379. CSLI Lecture Notes. → pages 163

Wooldridge, M. (2009). *An introduction to multiagent systems*. A John Wiley and Sons, Ltd. → pages 1

Young, H. P. (1975). Social choice scoring functions. *SIAM Journal on Applied Mathematics*, 28(4):824–838. → pages 87, 120

Young, H. P. and Levenglick, A. (1978). A consistent extension of Condorcet's election principle. *SIAM Journal on Applied Mathematics*, 35(2):285–300. → pages 95

Young, P. (1995). Optimal voting rules. *The Journal of Economic Perspectives*, 9(1):pp. 51–64. → pages 31

Zilberstein, S. (1998). Satisficing and bounded optimality. In *AAAI Spring Symposium on Satisficing Models*, Stanford, California. → pages 142, 182