Transformation of Propositional Attitudes in Epistemic Networks

Peter Baumann

München 2017
Transformation of Propositional Attitudes in Epistemic Networks

Peter Baumann
ABSTRACT

The analysis of group decisions is an active field of research whose importance increases due to higher connectivity in society. This paper analyzes deliberative and aggregative models of group decision making. Weaknesses and advantages of network epistemology and judgement aggregation are presented in detail. Furthermore combinations of both approaches are discussed since both models for themselves cannot account for real world group decision making accurately. In sum the paper tries to contribute to better understanding group communication, which in turn can be applied to various fields of research.

CONTENTS

1 Introduction 1

2 Propositional Attitudes 2

3 Aggregation vs. Deliberation 4

4 Network Epistemology 6
   4.1 Networks 6
   4.2 Optimality of Networks 7
      4.2.1 Information Direction and Costs 7
      4.2.2 Learning Problems 8
      4.2.3 Biased Agents 10

5 Theory of Judgement Aggregation 12
   5.1 Formal Model 12
   5.2 Conditions on Aggregation Rule 14
   5.3 Discursive Dilemma 15
   5.4 Escape Routes 17

6 Combined Models 19
   6.1 Probabilistic Aggregation and Richer Epistemology 19
   6.2 Judgement Transformation 20
   6.3 Strategic Manipulability 22
   6.4 Fair or Right Decisions? 22

7 Conclusion 24
1. Introduction

Communities of truth seeking individuals are omnipresent in modern society. In science, medicine, legislation, executive boards of companies etc. there are groups of people who presumably are interested in making the best decision given some evidence. Decisions in this sense may mean to decide whether or not a scientific theory is true, some patient has some disease, a defendant is guilty or an investment should be placed. Those groups are faced with the problem of determining a group decision given their individual beliefs or judgements. When analyzing such communities there are two different perspectives on how to look at the problem. First one can ask how the individuals within the group should communicate with each other to share their information and converge. Second one needs to analyze what the group's decision should be depending on the individual beliefs or judgements. Intriguing examples that illustrate the need for the scientific investigation of group decision are on the one hand the empirically observable asymmetry of the distribution of power in groups, on the other hand the inconsistencies which arise from majority voting.

This paper tries to give an introduction to both areas of research, which emerged from the two perspectives of "network epistemology" and "judgement aggregation". Firstly various aspects of how members of groups communicate with each other will be explored. Secondly problems of naturally motivated aggregation processes will be examined in detail. Thirdly some ideas and concepts on how to connect both approaches will be introduced and critically evaluated with respect to there applicability and accurateness.

The covered research fields gain importance due to a progressively connected world. Therefore a merged theory seems indispensable. Combined they can help understanding networks in various ways which is the key to optimizing the quality of beliefs or judgements of groups and therefore improving society in multiple ways.
2. Propositional Attitudes

When talking about group decisions in various contexts there are several terms to describe the attitudes individuals have. In different situations people hold different attitudes that need to be aggregated. While jurors hold judgements, doctors believe some patient has some disease and therefore judge a therapy to be the appropriate. Scientists may accept a certain theory or belief that a proposition is true and people in economical contexts often hold preferences. For a generalized theory of communication in epistemic networks it is important to clarify the terms and put them in connection with one another. While there is a whole area of research engaged with this issue, this paper will only point out the main differences and similarities of the concepts to demonstrate how they can be summarized into a more general concept.

All the above terms express some kind of attitude a subject has towards a proposition. The subject, e.g. the juror either thinks that a proposition p, say the defendant is guilty, is true or false. His attitude towards the proposition p can therefore be expressed as the assignment of a truth value to p. Similarly beliefs can be described. The main difference in this model is that a judgement is restricted to the value set \{t, f\}. In classical philosophy this holds for beliefs as well. A "belief" is "whenever we take something to be the case or regard it as true" [Schwitzgebel, 2015]. Since an agent can be more confident that a proposition p is true than he believes a proposition q to be true, degrees of belief were introduced. While degrees of belief can be modeled with many-valued logic, within this framework degrees of belief in the continuous interval between zero and one are allowed as well. A value of zero represents absolute refusal of a proposition whereas one corresponds to absolute certainty. Thus a belief can also be seen as a probabilistic judgement. It is clear that this is not at all a comprehensive description of the term belief and there might be a lot more differences between the terms "judgement" and "belief" in other dimensions of philosophical discourse.

A preference involves an ordering of two objects but it can be described as a judgement over the proposition "a is preferable over b". It can therefore be handled as a special case of judgements. To use a generalized concept [Dietrich and List, 2010] use the term "propositional at-
It describes the similarities of the terms with respect to the present discussion. Hereafter the term "propositional attitude" will be used to refer to the above terms. In addition the term "decision" needs to be clarified. Groups make decisions based on the individual attitudes. A decision will be thought of as an attitude a group holds toward a proposition. For example consider a group of doctors who have to decide whether or not a patient needs surgery. In this case "patient needs surgery" is the proposition, which the group either holds to be true or false. Their attitude of course may depend on the individual attitudes towards this and other propositions as well as some implications they hold to be true.
3. **Aggregation vs. Deliberation**

When a group of people who hold different propositional attitudes tries to come to a collective propositional attitude there are different possibilities how this transfer can happen. They can try to deliberate and ideally create consensus only by communication of information and evidence. The properties of this group communication model will be handled in the second chapter. On the opposite side they could simply hand in their set of propositional attitudes and then use an aggregation rule they previously agreed upon.

The deliberative model is insufficient in the sense that it is not at all clear why consensus could be attained in all cases. Additionally it is implausible to assume that it will be attained in most cases, since different inquirers can easily disagree on how to interpret data and therefore come to different judgements based on the same information. In these cases it is not guaranteed that talking about the issue will resolve the dissent.

The latter, static aggregation, does not run into the same kind of problems, since an aggregation rule does not need the agents to agree. An aggregation rule is constructed in a way that generates a group decision based on any individual opinions. However static aggregation quickly runs into inconsistencies if the process is required to have some very natural properties.

Therefore both models for themselves cannot provide a satisfying framework for group decision making. These circumstances give rise to some further questions. Firstly it needs to be examined if there is some combination of both models, which holds some new insight, e.g. some combination of deliberation and aggregation processes which have desirable properties and avoid inconsistencies or at least make them less likely. Secondly a natural question is to ask in which situations one model or the other or a combination is more suitable to come to a group decision.

This paper follows the structure of this chapter inasmuch as the next three chapters are concerned with deliberation, aggregation and combination respectively. In the fourth chapter a theory of information distribution and learning in groups will be introduced and applied to
several problems in social epistemology. The fourth chapter therefore deals with the deliberative aspect of group decisions. The subsequent chapter provides a formal model of judgement aggregation. Furthermore some natural conditions and resulting impossibilities in aggregation theory are discussed. The sixth chapter tries to compare both theories. In addition several combined models that try to include aggregation as well as deliberation are introduced and evaluated.
4. **Network Epistemology**

The traditional model of inquiry in epistemology features a single individual acquiring knowledge by looking at the world. The exchange of evidence and knowledge between different inquirers is either not part of the analysis or other inquirers and evidence drawn from the world are not handled separately. With increasing connectivity in epistemical communities greater attention must be paid to the epistemical analysis of groups of inquirers. Recent research in this field tries to analyze which properties of groups make them better epistemic groups. In this chapter an overview on the analysis of epistemic groups will be given. In the first section a mathematical model for group communication will be provided. The second section describes various aspects of optimal networks. It is split up in three parts. First general features of groups will be discussed. Second some specific learning situations will be covered. The third part deals with the problem of biased agents in groups.

4.1 **Networks**

From now on a group of inquirers will be referred to as a network. A network will be modeled as a mathematical graph with nodes corresponding to the group members and edges between them signifying some kind of information transmission. This way of modeling is taken from [Zollman, 2013](#). There can be different kinds of situations regarding the costs and direction of the information transmission. For instance the information can flow either both ways or just in one direction with one or both agents paying for the interaction. An asymmetrical distribution of costs or information transmission between the agents is modeled by a directed graph, i.e. the edges become arrows where information flows in direction of the arrows and/or the one at the end of the arrow pays the costs. Another way to characterize graphs is connectivity. A graph is connected if there exists at least one path between any two nodes. It is minimally connected if there exactly one path between two nodes. It is complete if there exists an edge between any two nodes. A directed graph is called strongly connected if there exists a directed
path between any two nodes, it is called weakly connected if the corresponding undirected graph is minimally connected. Within this framework the graph theory provides proofs for optimality in the different cases.

4.2 Optimality of Networks

In this chapter optimal networks for different communication situations will be introduced. The term optimality in this paper will be strictly utilitarian due to its compatibility to mathematical modeling. The first section is concerned with three general features of groups: information direction, costs and error possibility. In the second section some more specific situations are discussed. The third section introduces a theory to deal with biased agents in groups.

4.2.1 Information Direction and Costs

For the first part it is assumed that there is no error possibility, i.e. it is not possible that an information transmission can fail. First we analyze cases in which information flows both ways. If there are no costs involved any connected graph is optimal. Since the error possibility is zero they are all equivalent. If there are any costs involved only weakly connected graphs become optimal since they guarantee complete information distribution with minimal costs. In the case that information flows only one way the directed circle becomes the single optimal network since it is the strongly connected graph with the fewest edges [Zollman, 2013]. In the case where there is a slight probability that information transmission can fail, i.e. the quality of the received information depends on the distance between two nodes, only the "star" becomes optimal. A star is a graph that has a center node connected to all other nodes and which in turn are only connected to the center node. The different optimal graphs are shown in Figure 4.1.

In this sections closed static networks without new evidence and without any learning processes were discussed. In the subsequent section the model will be extended to more dynamic systems.
4.2.2 Learning Problems

This section will consider a variety of networks in more specific situations we different agents can "learn" from each other. Learning in this sense means that agents may change their attitudes based on the attitudes of the agents they get information from. So agents might hold a new propositional attitude given new information. There are several properties, which can help to characterize the different learning problems. Those properties are the following:

1. **Outside Evidence**
   If agents in a network get new information from outside the considered network during the learning process, it is said to have outside evidence.

2. **Independent Information**
   If the information the different agents have about the proposition are logically and
probabilistically independent, a network has independent information.

(3) **Second Hand Information**
A network has the "second hand information" property if any agent considers all other agents’ opinions independent of their distance in the network.

(4) **Independence of Propositions**
If the propositions the agents are sharing their information about are logically and probabilistically independent from one another, a network has the "Independence of Propositions" property.

In this chapter condition (4) will always be assumed, since this theory is not sufficient to deal with such cases. In chapter 6 condition (4) will no longer hold.

In the first part cases without outside evidence are discussed. Consider a situation where there is no second hand information but condition (2) holds. This means agents only consider the information of a smaller group of people in the community. The question then is weather a network of agents who hold different propositional attitude will converge to consensus given a certain communication structure. For example consider a group that needs to estimate a parameter \( p \). The communication structure could be that all agents average their estimate of the parameter over their opinion and the opinion of those who are connected to her, say her friends in the community. The key feature to optimality in these cases is regularity. Regularity means that each agent has the same amount of connections. These results hold as long as the group is concerned with propositions that satisfy condition (4) [Zollman, 2013]. Another interesting aspect is that weather an agent takes another agents opinion into account depends crucially on the opinion itself. In the above case where a parameter needs to be estimated, this could mean that an agent only average her estimate over agents whose estimate lies sufficiently close to hers. This tries to include the fact that an agent might distrust another agent’s information if it is fairly different to what she thinks is right. The model in section 4.2.3 tries, inter alia, to include this fact.

This paper will only cover one case where the group obtains new evidence over time. In the second part so-called bandit problems will be discussed. In bandit problems agents have to balance their interest in payoff and knowledge. The description follows [Zollman, 2013]. Consider a gambler who needs to choose between two slot machines. Both machines have
different payoff rates, which the agent does not know. The agent tries to maximize his payoff while he still wants to gather new information about the different machines. The agent can also consider independent information from other agents playing the machines. New information might lead to the insight that the other machine's payoff rate is probably higher and therefore rational to play. The optimal strategy for this problem is very complicated and therefore "bounded rational" strategies were introduced. These strategies are less complex and thus plausible for an agent to follow but they approximate the optimal strategy. Now consider two graphs. First an infinite line where all agents are only connected to their direct neighbours. Second add a single individual that is connected to all other individuals to the first graph. It can be shown that only the first, regular and minimally connected graph identifies the superior slot machine with certainty. Similarly in finite graphs minimally connected graphs become optimal. Both of these results are primarily due to the fact that more communication makes groups more vulnerable to statistic uncertainty [Zollman, 2013].

4.2.3 Biased Agents

For the previous chapters it was assumed that all agents within the network are honest truth seekers. However there are cases where agents are biased and follow other strategies than approaching the truth best to their knowledge. This could be for example an agent whose primary goal is to maximize profit. For example a pharmaceutical company, a lobbying agent on a company board or funded research in science might be subject to this problem. Biased agents can prevent the group from converging to the truth. Consider a bandit problem with two machines A and B with payoff rates 1.01 and .98 respectively. The agents try to figure out with machine has the better payoff rate. Assume all agents start with the true payoff rates. This is no restriction, since this is the scenario most likely to converge. If it does not converge no other will. The agents will start using machine A and only swap if by chance the acquired evidence about machine A will drop under .98. This system will converge and in the long run all agents will only use machine A.

Now a biased agent is introduced, assume the bias is .05 in favor of machine B. In the case that all agents use machine A, the only evidence they gather about machine B is from the biased agent who reports a payoff rate of 1.03. Therefore some agent will eventually start using machine B. Then evidence about machine B is obtained by this agent plus the biased agent. Averaging their evidence gives a reported payoff rate of 1.05 which in turn will make agents swap back to machine A until all agents but the biased agents use machine A. This process recurs infinitely preventing the network from converging. The example and the result
are taken from [Holman and Bruner, 2015]. This happens mostly due to the static structure of communication. Agents cannot choose whom to talk to and they weigh the evidence of all other agents equally. [Holman and Bruner, 2015] suggest a different communication practices to avoid this result. The so-called "Endogenous Network Formation". Roughly in this dynamic network an agent's weight on a specific agent's opinion depends on the difference between her and the other agents opinion in the previous rounds. This procedure makes the agents identify the biased agent and discard his opinion in the long run, i.e. the weight of the biased agents opinion converges to zero.
5. **Theory of Judgement Aggregation**

The theory of judgement aggregation tries to formalize the act of bringing together individual judgements into a group judgement. There are some very natural requirements this act should meet. The interesting result that those requirements cannot be met at the same time is called the discursive dilemma. It is not limited to judgements. The aggregation of most propositional attitudes, e.g. preferences, probabilities or utilities are subject to the discursive dilemma. Since then a tremendous amount of research was concerned with further understanding and trying to dissolve the dilemma. This chapter introduces the theory of judgement aggregation in four parts. In the first part a formal model of judgement aggregation is explained. The second part lists some desirable properties of an aggregation rule. The third part is concerned with the discursive dilemma. Ultimately some properties are inspected more thoroughly in order to provide some possible strategies to avoid the dilemma.

### 5.1 Formal Model

To discuss impossibility results and to see how different properties of the procedure affect the aggregation, it is necessary to provide a formalization of the aggregation of judgements. The model will be introduced for propositional attitudes in general, since judgements are just a special case of propositional attitudes and the general model is not more difficult but does hold some further insights. The formal model for the propositional attitude aggregation follows [Dietrich and List, 2010].

Consider a situation where a number of members of a group N all have some individual attitudes towards a set of propositions X, which is closed under negation, i.e. \( p \in X \Rightarrow \neg p \in X \). X is also called an agenda. An attitude is represented by an attitude function \( F : X \rightarrow V \). \( V \) is the set of possible values. \( V \) may be the interval \([0,1]\) for probabilities or \([0,1]\) for judgements representing truth and falsehood of propositions. The individual attitude functions need to be rational. An attitude function is rational if it is consistent, i.e. if it is extendable to
a well-defined valuation function, e.g. for a value set \( \{0,1\} \) the logic closure of the set \( F^{-1}(1) \) and \( F^{-1}(0) \) need to be logically consistent. Let \( T \subset V^X \) be the subset of consistent attitude functions where \( V^X \) denotes the space of functions from \( X \) to \( V \). A set of \( N \) rational attitude functions is called a profile \( P \). \( P \) is an element of the \( N \)-fold Cartesian Product of \( T \). An aggregation rule or function \( A \) needs to find an attitude function given a certain profile, i.e. it needs to map \( P \) onto a function \( F \). In short:

\[
A: T^N \rightarrow V^X
\]  

(5.1)

To illustrate the model an example is provided. Consider a group of three judges A, B and C that have to come to a joint decision about the following propositions:

p: "The defendant had motive for the murder"
q: "The defendant had opportunity to commit the murder"

The agenda \( X \) in this case is \( X = \{p, \neg p, q, \neg q\} \). Since the judges can only judge a proposition to be true or false, the value set is given by \( V = \{0, 1\} \). Assume that A thinks that the defendant had motive but no opportunity while B judges that he had no motive but did have the opportunity. Judge C is convinced of both \( p \) and \( q \). The individual attitude functions are therefore given as:

\[
F_A: X \rightarrow V, x \mapsto \begin{cases} 
1, & x = p, \neg q \\
0, & x = q, \neg p
\end{cases}
\]  

(5.2)

\[
F_B: X \rightarrow V, x \mapsto \begin{cases} 
1, & x = q, \neg p \\
0, & x = p, \neg q
\end{cases}
\]  

(5.3)

\[
F_C: X \rightarrow V, x \mapsto \begin{cases} 
1, & x = q, p \\
0, & x = \neg p, \neg q
\end{cases}
\]  

(5.4)

It is clear that all those functions are elements of \( Q \). Two examples for aggregation rules in this example are:

\[
A_1: T^3 \rightarrow V^X, t = (t_1, t_2, t_3) \mapsto t_1
\]  

(5.5)

The aggregated attitude function in this case is \( F_A \) independent of the other input functions.

\[
A_2: T^3 \rightarrow V^X, t = (t_1, t_2, t_3) \mapsto t_d = [(t_1 + t_2 + t_3)/2]
\]  

(5.6)

This rule represents majority voting on the propositions.
5.2 Conditions on Aggregation Rule

This section will introduce five conditions on the aggregation rule as a function, which can be motivated naturally and seem plausible to request from an aggregation function. These conditions can be traced back to majority voting and democratic practices as well as the normal usage of the terms "aggregation" and "judgement". The definitions of the conditions follow [List, 2012]. The conditions are as follows:

**Universal Domain**

The Universal Domain condition (UD) states that any profile is admissible to the aggregation function. So any set of rational attitude functions is allowed as input to the aggregation function. More precisely this means that any agent should be allowed to have any kind of attitude function independent of any other attitude function as long as it is consistent. Technically the Domain of $A$ is the whole Space $T^N$.

**Collective Rationality**

The condition of Collective Rationality (CR) restricts the output of the aggregation function to such attitude functions, which are rational as well as complete. A collective propositional attitude function is rational if it is consistent as defined above, i.e. $A$ maps onto $T$. The completeness states that the aggregated attitude function assigns a value to all propositions in the agenda, i.e. the range is restricted to surjective $F \in T$.

**Independence / Systemacy**

The value the aggregated function assigns to a proposition should be independent of the values the attitude functions in the profile assign to other propositions. This constraint is called Independence (IND). If in addition the aggregation rule is the same for all propositions, it is systematic (SYS).

**Monotonicity**

Consider a given profile. Given that profile the aggregation function ascribes a value $t$ to some proposition $p$. If now an attitude function which ascribed a value $s \neq t$ to proposition $p$ is swapped for another attitude function which does ascribe $t$ to $p$, then the aggregation function still ascribes $t$ to $p$. In other words: If one attitude function gets closer to the aggregated attitude function, ceteris paribus, then the new aggre-
5.3 Discursive Dilemma

gated attitude function remains the same. In the following the monotonicity condition will be referred to as (MON).

**Unanimity Preservation**

If a profile consists of N equal attitude functions then the aggregated attitude function is exactly the same as the individual ones. Unanimity Preservation (UP) is intuitive, since a unanimous profile is best represented by the individual attitude function it contains.

### 5.3 Discursive Dilemma

For the restricted binary value set \{0, 1\} [List, 2012] proofed that only dictatorships meet all the above conditions at the same time. A dictatorship is an aggregation rule such that there exists an index n such that for all propositions in X it holds that the aggregated function is equal to the attitude function \( F_n \) of some fixed individual. Formula (5.5) is an example for such a rule. Note that in literature the dilemma is sometimes stated differently. A sixth condition, namely Non-dictatorship, is added. This sixth condition leaves no possible attitude function.

Another implicit condition that needs to be mentioned is that the agenda needs to be sufficiently complex. If (SYS) is assumed, it needs to contain a minimally inconsistent subset. A minimally inconsistent subset is a set, which is logically inconsistent but every proper subset is consistent. If the agenda in section 5.1 is expanded by the propositions r: "The defendant is guilty" and s: "If a defendant had motive and opportunity, he is guilty" as well as its negations, a minimally inconsistent subset can be constructed. Consider \( I = \{p, q, \neg r, s\} \subset X \), with \( s = ((p \land q) \rightarrow r) \). I is obviously inconsistent but each proper subset of I is consistent.

If only (IND) is assumed the agenda needs to be totally-blocked. An agenda is totally-blocked if for every two propositions p and q in the agenda X there exists a subset R of X which contains p and logically implies q. The agenda in the example is not totally-blocked, since for example there exists no subset that contains p and logically implies s.

Amongst others [List and Pettit, 2002] called the result the discursive dilemma, because it shows that there is no simply methodology to merge individual binary attitudes to a collective one.

[Dietrich and List, 2010] used the same model as in this paper to show a generalized result. In
5.3 Discursive Dilemma

cases with non-binary value sets, weighted linear averaging rules satisfy the conditions.

\[ A = w_1 \cdot F_1 + w_2 \cdot F_2 + w_3 \cdot F_3, \quad \text{with} \quad \sum_{n=1}^{3} w_n = 1 \]  \hspace{1cm} (5.7)

In the case of binary value sets there is only one weighted linear averaging rule that guarantees that the rule is well defined – the dictatorial rule. Well defined means that the function (5.1) maps onto \( V^X \). That means the function \( F \) needs to map onto the binary Value set \( V \). In other words the weighted average over the individual attitude functions is only necessarily a function with range \( \{0, 1\} \) if all weight is given to one individual.

To give an example of an inconsistent aggregation rule consider the judges in section 5.1 with the agenda expanded as above. The judges hold the same judgements as before but in addition they all grant proposition \( s \), since it is assumed this rule is given by the law. Since the individual attitudes need to be consistent judge C deems the defendant guilty while A and B judge the opposite. The following table summarizes the judgements.

<table>
<thead>
<tr>
<th></th>
<th>( p )</th>
<th>( q )</th>
<th>( s )</th>
<th>( r )</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>✓</td>
<td>X</td>
<td>✓</td>
<td>X</td>
</tr>
<tr>
<td>B</td>
<td>X</td>
<td>✓</td>
<td>✓</td>
<td>X</td>
</tr>
<tr>
<td>C</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 5.1: The table shows the Judges’ attitudes towards the four propositions in the example.

Consider the majority voting rule (5.6). The aggregated attitude function is:

\[ t_d(x) = \lfloor (t_1(x) + t_2(x) + t_3(x))/2 \rfloor \quad x \in X \]  \hspace{1cm} (5.8)

For \( x = p \):

\[ t_d(p) = \lfloor (t_1(p) + t_2(p) + t_3(p))/2 \rfloor = \lfloor (1 + 0 + 1)/2 \rfloor = 1 \]  \hspace{1cm} (5.9)

The following table shows the aggregated values assigned to the propositions. Now if we consider the set \( t_d^{-1}(1) \) of judgements the group grants using this rule, we obtain the logically

\begin{array}{cccc}
  t_d(p) & = & 1 & \quad t_d(q) = 1 & \quad t_d(r) = 0 & \quad t_d(s) = 1 \\
  t_d(\neg p) & = & 0 & \quad t_d(\neg q) & = & 0 & \quad t_d(\neg r) = 1 & \quad t_d(\neg s) = 0 \\
\end{array}

Table 5.2: The table shows the output of the aggregated attitude function \( t_d \).
inconsistent set \( \{p, q, \neg r, s\} \). So in this rather simple example the (CR) condition is violated by majority voting.

The result is a generalization of Condorcet's Paradox, which states that a group of transitive preferences might lead to a cycle collective preference. Arrow's theorem states a similar impossibility result in the case of preferences, which can be modeled as attitudes in the above framework.

Since the above conditions are logically tight, i.e. they are sufficient and necessary for the result one might try to relax one or more of the conditions in order to prevent a dictatorship.

In the following section several escape routes will be discussed.

### 5.4 Escape Routes

To avoid the discursive dilemma one needs to leave the logic space outlined by the five conditions in the previous section. The CR condition is hard to give up since the consistency of the output is crucial. The output is said to be a collective judgement or preference or any other kind of attitude and the terms "judgement" or "preference" already contain a notion of consistency. The completeness as pointed is important for practical reasons. Theoretically a collective judgement could just be an attitude function restricted to a subset of the agenda such that the function is consistent. Practically it is desirable to have the group hold an attitude towards all propositions of the agenda, since in most cases they have to make a decision and cannot just be undecided.

UP and MON are conditions, which are motivated by the intuition of majority voting. They seem closely related to the term "aggregation". "Aggregating" as in "bringing together" or "accumulate" already includes some notion of continuity and information preservation, i.e. an "aggregation" is a process where information is condensed and represented. On the other hand a process whose result is not closely related to the input information would not be called an "aggregation". Looking closer at the matter it is not clear why UP needs to hold. As can be seen in section 6.1, in a richer epistemological model UP might not be acceptable.

The last two conditions UD and IND seem to be candidates for relaxation. This part will give a first introduction to the concepts that discard one of both conditions. First the relaxation of UD will be examined, subsequently IND/SYS is abolished.

So what does it mean that UD is violated? From an individuals perspective it seems indispensable. Generally an individual wants to be free to hold any attitude function as long as it
is consistent. More specifically consider an individual confronted with a set of other attitude functions from the rest of the group, which then might restrict his freedom to hold any attitude – possibly to the point where he has no choice left. Then he was arguable not even part of the aggregation process. But from a more holistic view one might try to create some kind of "consensus" before the aggregation such that the domain is restricted to a sufficiently small set in order to prevent any inconsistencies. An important example for a restricted domain that guarantees consistency with majority voting is unidimensional alignment. A profile is unidimensionally aligned if it is possible to find an order for all N group members such that for all propositions there exists an index n such that the agents one to n as well as the agents n+1 to N hold the same judgements. To illustrate this consider table 5.1. For the propositions q, r and s n is 1, 3 and 2 respectively but for p there is no such n. In fact there is no other ordering of the judges A, B and C that provides an unidimensional alignment. If judge C grants ~p instead of p, table 5.1 would constitute such an alignment and at the same time the majority voting would become consistent. In section 6.2 possible network communication strategies that might provide this "consensus" will be discussed.

Only relaxing SYS does not seem sufficient since with little more complex agendas the impossibility still holds. Therefore the next part explores cases where IND is violated. Giving up independence means that attitudes an agent and the group hold towards the different propositions are not independent from one another. There exist different aggregation rules where this condition is violated. First consider a ranking over all propositions in the agenda. Then let the attitudes be aggregated propositionwise corresponding to their rank in importance. The group starts with the most important proposition and takes a vote on that proposition. The result will be included in the group attitude function. If later in the process the result of a vote would make the group attitude function inconsistent it will be ignored and the negation will be included instead. This way the aggregated function stays consistent and the more important a proposition is, the more likely its group attitude is represented by the majority vote. A second way to aggregate without independence will be introduce in section 6.2.

The problem that comes with the relaxation of IND is, that it is a necessary condition for non-manipulability. Strategic manipulability describes that an agent might be able to hold an inaccurate attitude function in order to make his true attitudes more likely to become the aggregated attitude. The strategic manipulability can be countered in specific frameworks one of which is discussed in section 6.3.
6. Combined Models

For the combination of the models it is not sufficient to only differentiate between deliberation and aggregation. There is a distinction that some philosophers tried to bring up that is important for aggregation in networks. [Pivato, 2008] and [Williamson, 2009] argue that there is a difference to be made whether a group is concerned with a "fair" or a "right" aggregation. While a lot of research in judgement aggregation focuses on fair aggregation, partly represented by the conditions in chapter 5, the results do not necessarily hold if one is concerned with right judgements. A fair aggregation is one that equally accounts for each individual's opinion while a right aggregation is one that represents the world best. This chapter provides some approaches on right as well as fair aggregation processes that include deliberation.

6.1 Probabilistic Aggregation and Richer Epistemology

This section introduces two models that focus on a right rather than a fair aggregation. Pivato uses statistic opinion pooling while Williamson refers to a richer epistemology. These examples develop from the theory in chapter 4 but with the important difference that propositions are not independent but sufficiently complex entangled. Pivato argues that a binary value set is inadequate since two agents cannot both believe with certainty some contradicting propositions. More precise any T-valued logic is insufficient to represent beliefs held over propositions. He argues that the subjective probability estimates over propositions, beliefs in continuous value set strictly between zero and one, need to be aggregated. This transfers the problem to the problem of statistic opinion pooling, which is a research area of its own. The problem is that this leads to similar results, i.e. probabilistic versions of the discursive dilemma. Therefore he argues that this simple aggregation is underspecified. Additional data on the personal information of the agents, i.e. how the judgements are justified needs to be added. With full disclosure of personal data everyone gets all the personal information of everyone else. This is not a very realistic scenario as Pivato realizes...
Judgement Transformation

Judgement transformation stays within the framework of chapter 5 and is more concerned with deliberation and aggregation to create a fair outcome. Judgement transformation tries to integrate pre-decision communication and into the model of chapter 5. This might lead to some restriction of the universal domain condition, which in turn might avoid the discursive dilemma.

The theory of judgement transformation was introduced by [List, 2011]. In contrast to an aggregation function a judgement transformation function maps profiles onto profiles, i.e. the function has a set of attitude functions as input as well as output. The new set of attitude function can be seen as the post communication attitudes the agents hold. Lists shows that under some similar conditions this transformation function runs into a dilemma like aggregation
functions do. In fact the only transformation function that satisfies those similar conditions is the identity function, i.e. nobody changes their mind.

He argues that the goal of pre-decision communication cannot be consensus, since this would make the problem collapse into the judgement aggregation problem. The transformation process might lead to kind of "meta agreement" that in turn restricts the domain for the subsequent aggregation process sufficiently. As seen before single-peakedness is sufficient for consistent majority voting. Therefore List tries to come up with a transformation process that provides single-peaked profiles as output.

To make effective communication possible he relaxes the IND condition. He introduces a new class of transformation function, which are not some kind of sequential priority rules that were presented in section 5.4. The constrained minimal revision functions (CMRF) work in two steps. In the first step a set of admissible output attitude functions is selected. This can be thought of as a focusing function. In a second step each individual chooses the attitude function closest to her pre-communication attitude function. Closest means in terms of some metric on the attitude functions. List proposes a focusing function for preferences that works with single peaked preferences – a condition similar to unidimensionally aligned judgements but for preferences. A natural metric on the attitude functions is the Hamming metric. It counts the propositions for which the attitude functions give a different output. This transformation guarantees single peaked profiles such that in a next step a consistent majority vote can be taken.

Since List only introduces a CMRF for preferences, in this section a focusing functions for judgements will be suggested. Consider the profile $X$ consisting of $N$ attitude functions. In a first step a maximal unidimensional alignable subset $U \subseteq X$ of the profile is identified, i.e. the subset that contains the most attitude functions but is still unidimensionally aligned. Now an admissible functions $F$ such that $F \cup U$ can still be unidimensionally aligned. This defines a focusing function. In the second step each individual needs to select a post communication attitude function from the set of admissible functions. This happens similarly to List’s second step. Each individual identifies the admissible attitude function closest to his initial function by calculating the Hamming distances. The function with the smallest Hamming distance to the initial function is the agents new attitude function. This procedure guarantees unidimensionally aligned profiles in the case of judgement aggregation and is therefore applicable to majority voting.
6.3 Strategic Manipulability

Nonetheless one deficiency remains in the model of the previous section. The procedure is susceptible to strategic manipulation. This is a serious problem which can be countered in two ways. On the one hand one can try to work out what makes group members more honest. For instance some kind of moral obligation to answer truthfully. Though this is an interesting philosophical question it cannot be answered within this paper. On the other hand and more relevant for the present topic one can try to find a mechanism that tries to expose biased agents. A first idea is to transfer [Holman and Bruner, 2015] model that decreases the influence of biased agents progressively to [List, 2011] CMRF. If one interprets the judgement transformation as rounds of deliberation it might be possible to identify the manipulative group member. The following is an outline of how this could be implemented. The manipulator needs to calculate a few rounds in advance to know what is best attitude function to hold. Assume the agent cannot calculate all the way to the end of the process. He might then have to change her attitude function drastically at some point in the process due to new calculations. This is where one could try to intervene since all honest agents try to stay as close to their initial attitude as possible. Imagine one agent is suspected of manipulation due to drastically changing attitude functions. How can his influence be lowered? All agents chose their transformed attitude function by calculation the closest admissible function. In List's model this is calculated using the Hamming metric. Therefore some kind of weighed Hamming metric needs to be introduced that includes not only proximity to ones initial attitude but also difference from the biased agents attitude. This could be implemented similarly to round based updates of weighs for network communication in [Holman and Bruner, 2015].

This paper only gives an idea of how strategic manipulability can be avoided in the case of CMRF. A thorough study of this matter can be an interesting topic for further research although one should hold in mind that manipulability cannot be ruled out totally as long as the process is theoretically calculable. But since group decision making is mostly done by humans with limited computing capacity, this might not be a problem.

6.4 Fair or Right Decisions?

As can be seen in the above approaches, deliberation is the key ingredient for avoiding the discursive dilemma in both fair an right aggregation models. The line between right and fair cannot be drawn as clearly as Pivato and Williamson claim.
For one thing a right aggregation can be wrong and vice versa. Based on the set of information a group can come to an aggregated judgement which turns out to be wrong afterwards. So aggregation is focused on whether a judgement is right based on the information within the group before the process. Secondly in real world group deliberation it might not be clear what is evidence, what is belief and what is judgement, e.g. the members might not agree upon what counts as evidence. Thirdly if the group shares information and evidence then there might not be an agreement on how to interpret this evidence or information. So unless they agree upon the interpretation it is not clear which is the "right" judgement based on the evidence or information. In this case a "fair" judgement might be the only "right" one in the sense that everyone's interpretation of the evidence counts. The assumption that they all agree upon the interpretation is not plausible aside from some very basic facts maybe. This variance in interpretation is what defines a group of individuals. Otherwise the group would collapse to a single individual with an evidence set and some interpretation function.

To sum up there is clearly a difference in the terms "right" and "fair" in various ways, but when it comes to group decision making they cannot be completely separated from one another. Therefore models which only consider one of both aspects are not sufficient in describing real world decision processes.
7. Conclusion

This paper presented a detailed introduction to the theories of network epistemology and judgement aggregation. Propositional attitudes were introduced, situations in which groups were confronted with learning processes were examined and a formal model of aggregation of attitudes was provided. Furthermore some impossibility results in aggregation theory were introduced and several routes to escape these impossibilities were thoroughly investigated. Ultimately the theories were extended to models that try to include deliberation as well as aggregation. On the one hand these models helped to carve out the importance of deliberation in the aggregation process. On the other hand they helped to conclude that in most cases not gainful to consider the aim of a decision, i.e. distinguish between fair and right decisions.

A critical question that remains is weather the provided theories can model real decision making accurately. The increasing complexity of models that try to include more and more aspects of decision making run the risk of moving away from real life decision practices. For instance it is not clear if group members act rational and what other aspects away from objective theory influence the groups’ decision. The models in this paper fail to include psychological effects and more generally the humaneness of group members. Especially when it comes to communication this aspect should not be omitted.

Nonetheless the aggregation of attitudes remains a problem that lots of people and companies are confronted with on a daily basis. Though more research needs to be done to further understand group decision making this paper tried to contribute to solving the decision problems that doctors, judges, executives and researchers are faced with. There are interesting routes that need to be explored in further studies. It would be desirable to provide a complete model of transformation of propositional attitudes as well as a precise examination of strategic manipulability that includes designs and moral aspects as well as formal modeling. Furthermore a detailed philosophical discussion of the differences and similarities of right and fair decisions is an interesting possibility for further research.
BIBLIOGRAPHY


LIST OF FIGURES

4.1 Optimal Networks ......................................................... 8
List of Tables

5.1 Judges' Attitudes ................................................................. 16
5.2 Aggregated Attitude ............................................................. 16
EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Arbeit ohne Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher in gleicher oder ähnlicher Form in keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.

Ort, Datum

Unterschrift