Propositional Opinion Diffusion with Constraints

MSc Thesis *(Afstudeerscriptie)*

written by

Sirin Botan
(born July 31st, 1991 in Oslo, Norway)

under the supervision of **Umberto Grandi** and **Ulle Endriss**, and submitted to the Board of Examiners in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam.*

Date of the public defense:	**Members of the Thesis Committee:**
September 26, 2016	Prof Jan van Eijck
	Dr Ulle Endriss
	Dr Umberto Grandi
	Prof Sonja Smets
	Dr Jakub Szymanik (chair)

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

Abstract

This thesis examines influence-based iterative opinion diffusion mechanisms on social networks where agents update their opinions based on the opinions of their influencers in the social network. We use tools from judgment aggregation to analyse the strengths and weaknesses of an existing opinion diffusion framework, Propositional Opinion Diffusion, and define a variant of this type of diffusion mechanism, Propositionwise Opinion Diffusion, which guarantees the rationality of agents' opinions at every step of the diffusion process. We compare the two variants of the mechanism, both by using an axiomatic approach, and by exploring when they might coincide or differ on the opinions agents hold when the mechanisms terminate.

Acknowledgements

Though I was not able to list all three on the title page, this thesis was written under the supervision of Umberto Grandi, Ulle Endriss and Laurent Perrussel.

Laurent was invaluable to me as a supervisor. I would like to profoundly thank him for all his help and advice, and for welcoming me to Toulouse with open arms. I am beyond grateful I was given the chance to meet such a kind, funny and encouraging person. This thesis would certainly not have been possible without him.

To Umberto, the second member of my Toulouse welcome committee, I also extend my deepest, most sincere gratitude for his guidance and patience through countless drafts. He has supported me through both fun times and frustrating ones. I consider myself very lucky I got the chance to learn from Umberto; it was a true joy to be his student.

Finally, I have to thank my “auxiliary supervisor” Ulle, without whom I would never have gotten the chance to travel to Toulouse to write my thesis. I owe so much to Ulle, who first introduced me to the exciting world of computational social choice and has always been a source of encouragement and support. I could not ask for a nicer person to have in my corner.

Contents

1 | Introduction and Motivation

Judgment Aggregation [List and Pettit, 2002] studies how to take the opinions of individual agents about some given set of issues or propositions and aggregate them to find the group's opinion, possibly in order to make some collective decision. Take the following example: a hiring committee needs to decide if a certain candidate is qualified for a position. There are three people on the committee and hiring decisions are made based on the majority opinion. Suppose each committee member needs to decide whether the candidate has a suitable education (p) and whether the candidate has an appropriate amount of experience (q). The only requirement for the committee members is that if they are of the opinion that both p and q, then they must also want the candidate to be hired for the position (r) – i.e. if a candidate both has a suitable education and an appropriate amount of experience, she should be hired. In other words, there is a constraint, $(p \wedge q) \rightarrow r$, which all the agents (or committee members) are required to satisfy. We can represent this in the following table where a 0 means an agent rejects the proposition (i.e. accepts the negation), and a 1 means she accepts it.

	p	q	r
Agent 1	0	1	0
Agent 2	1	0	0
Agent 3	1	1	1
Majority	1	1	0

Notice that despite the fact that each individual agent satisfies the given constraint, taking the proposition-wise majority results in a collective opinion that violates it – both p and q are accepted by a majority of agents, while r is not. This is an example of the *discursive dilemma*, and is a commonly encountered problem in judgment aggregation [List and Puppe, 2009]. One possible solution to counter such irrational collective opinions is to introduce interaction or communication between agents into the decision making process; this is the approach we will examine in this thesis.

What may happen if the committee members deliberate about each candidate before making a collective decision? This deliberation process can lead to a change in the individuals' opinions – and as a consequence the collective opinion – depending on the dynamics of the group of agents. Perhaps there is one senior member who influences all others' opinions to change, perhaps there are pairs of committee members who influence each other, or a cycle of influence between all three. Based on this influence relation and everyone's initial judgments, the opinions of agents might change several times during deliberation before they reach a stable point at which the committee can make a final decision. This process of changing opinions based on social influence is what *Propositional Opinion Diffusion* (POD) attempts to formalise.

In real life examples of opinion diffusion and social influence, it is rarely the case that all agents in a network are influenced by – or even able to know – all other agents' opinions. This is especially true if we have a large number of agents. To reflect this POD employs a social network which limits communication between agents and designates which agents are connected and thereby able to influence one another. Thus, the particular type of communication we model does not assume that a global discussion between all agents is always possible. In this way, the structure of the social network is allowed to play a significant role in the opinions that agents hold post-communication.

Models of social influence have been studied from many angles. DeGroot [1974] and Werner and Wagner [1981] study opinion diffusion as a means of reaching consensus within a group. Yet others, such as Miller [1992], Knight and Johnson [1994] and Dryzek and List [2003], use the notion of inter-agent communication as a means of reconciling ideas from *deliberation theory* with those from *social choice* and judgment aggregation. *Belief merging* [Konieczny and Pino Perez, 2002], provides an alternative approach to modeling social influence on a network. Schwind et al. [2015] use *belief revision games* as a way of modeling how beliefs propagate in a social network of agents, and Schwind et al. [2016] examine whether promoting a belief will always lead to wider acceptance of the belief by agents in the network. We can also find social networks playing a role in other game-theoretic studies of opinion diffusion. Simon and Apt [2015] use threshold models on social networks where agents choose to adopt certain products based on the ones adopted by the agents they are connected to in the network, and show that on certain networks a product can be adopted by the whole network. Apt and Markakis [2011] take these threshold models a step further and explore the strategic interplay between agents in such a network. Diffusion on social networks has also been studied in logic by Christoff and Hansen [2015] who developed a dynamic modal logic to reason about the diffusion of not just opinions but products, behaviors and even diseases on a social netowork. Baltag et al. [2016] studied the epistemic aspects of threshold models of social influence and explored these in a logical framework. Here, we study a model of opinion diffusion based on social influence using tools from judgment aggregation [List, 2012]. The starting point of our exploration is work by Grandi et al. [2015] and Brill et al. [2016], who defined models of opinion diffusion on binary issues and preference diffusion, respectively. Both use a social network as a central part of the diffusion mechanism.

This thesis aims to expand upon work already done on models of influence and the diffusion of opinions in a multi-agent system. The framework of Propositional Opinion Diffusion which is the basis of our work was introduced by Grandi et al. [2015]. While their framework allowed for opinions on multiple issues or propositions, the aggregation of opinions on multiple issues – with the constraints that accompany them – was not extensively explored. This will be the focus of this thesis. We aim to explore what type of decision a group will land on following an opinion diffusion process based on social influence, as well as how we can avoid agents being influenced to change from a rational opinion to an irrational one.

1.1 Binary Aggregation with Integrity Constraints

To model opinion diffusion on a network, we will be using the framework of Binary Aggregation with Integrity Constraints [Dokow and Holzman, 2010a], [Grandi and Endriss, 2011]. We will see that the we can model judgment aggregation problems – such as the example at the beginning of this chapter – in the framework of binary aggregation. The type of aggregation problems we will examine will be those where agents are asked to give a yes or no answer over a set of related issues where a *yes* is encoded as a 1 and a *no* as a 0. When there are multiple issues over which agents must give opinions, we assume that these issues are interrelated in some way.

The formal framework consists of a finite set of m propositional variables or *issues*, $\mathcal{I} = \{p_1, \dots, p_m\}$, where each issue represents a binary choice. We call $\mathcal{D} = \{0,1\}^{\mathcal{I}}$ the domain associated with this set of issues. For a finite set of agents, $\mathcal{N} = \{1, \dots, n\}$ we call $B_i \in \mathcal{D}$ the opinion or *ballot* of agent $i \in \mathcal{N}$ over each of the issues in $\mathcal{I}$. A vector of all the ballots of agents in $\mathcal{N}$, $\boldsymbol{B} \in \mathcal{D}^{\mathcal{N}}$ is a *profile*. $B_i(p)$ is agent i's judgment on $p \in \mathcal{I}$ in the profile $\boldsymbol{B}$. We write $\boldsymbol{B} =_{-i} \boldsymbol{B'}$ to mean two profiles $\boldsymbol{B}$ and $\boldsymbol{B'}$ are identical if we ignore agent i's ballot, and $\boldsymbol{B} =_{-p} \boldsymbol{B'}$ to mean two profiles $\boldsymbol{B}$ and $\boldsymbol{B'}$ are equal when we ignore the judgments on $p \in \mathcal{I}$. For a profile $\boldsymbol{B}$ and an issue $p \in \mathcal{I}$, $N_p^{\boldsymbol{B}} = \{i \in \mathcal{N} \mid b_i(p) = 1\}$ is the set of agents who support p and $\overline{N_p^{\boldsymbol{B}}} = \{i \in \mathcal{N} \mid b_i(p) = 0\}$ is the set of agents who do not support p. We say a profile is *unanimous* if every ballot in the profile is the same.

Let $\mathcal{L}_{\mathcal{I}}$ be the propositional language corresponding to the set of issues $\mathcal{I}$, meaning $\mathcal{L}_{\mathcal{I}}$ is the set of propositional formulas which use only the propositions in $\mathcal{I}$. For example, if $\mathcal{I} = \{p, q, r\}$ then p, $(p \wedge r)$, $(q \rightarrow p)$ and $(p \wedge r \leftrightarrow q)$ are all formulas of $\mathcal{L}_{\mathcal{I}}$. For any $\varphi \in \mathcal{L}_{\mathcal{I}}$, let $\mathrm{Mod}(\varphi)$ be the set of models that satisfy φ. An *integrity constraint* is formula $\mathrm{IC} \in \mathcal{L}_{\mathcal{I}}$. Any integrity constraint defines a domain of aggregation, $\mathrm{Mod}(\mathrm{IC})$, of the ballots (or models) which satisfy the integrity constraint. We can view these integrity constraints as rationality requirements for the agents' opinions, given the issues in $\mathcal{I}$ and the problem at hand.

An example of a rationality requirement in judgment aggregation is that an agent which accepts a formula p must reject the formula $\neg p$. In preference aggregation a common rationality requirement is that each agent has transitive preferences, meaning if she prefers an alternative a to alternative b, and prefers b to a third alternative c, then she must also prefer a to c. Such rationality requirements can be expressed with a formula – the integrity constraint – ensuring that the agents which satisfy the formula are rational with regard to the interdependencies of the propositions in $\mathcal{I}$. These examples do not take into account what the propositions mean for a given aggregation problem, but we can also use integrity constraints to impose rationality requirements which are specific to a concrete problem. Recall the example of the committee at the beginning of this section. There, we required each agent to accept $(p \wedge q \rightarrow r)$, though this was not because of any logical connection between the propositions, but rather a requirement for an opinion to be feasible for the given problem and the given meaning of p, q, and r.

The benefit of representing the rationality requirements of agents as propositional formulas in the form of an integrity constraint is that we can analyse the syntactic properties of constraints which behave in certain desirable ways, or gain some insight into the common syntactic properties of constraints which we find reasonable or useful. We find one example of this in the work of Grandi and Endriss [2011], who characterize the class of integrity constraints that guarantee that the outcome of a certain aggregation rule will be rational (i.e. will satisfy the constraint), if each individual agent submits a rational opinion.

Formally, a *binary aggregation problem* $\mathcal{J} = \langle \mathcal{I}, \mathcal{N}, IC \rangle$ is defined by a set of issues $\mathcal{I}$, a set of agents $\mathcal{N}$ and an integrity constraint IC. Thus, a binary aggregation problem defines the space of possible collective decisions over the set of issues $\mathcal{I}$ for the agents in $\mathcal{N}$. Given a set of issues $\mathcal{I}$ and a set of agent $\mathcal{N}$, an *aggregation rule* F is defined as a function

$$F : \mathcal{D}^{\mathcal{N}} \rightarrow \mathcal{D}$$

which maps each profile of ballots to an element in the domain (i.e. a ballot). Most aggregation rules are defined for any possible profile of rational ballots – when this is the case we say that the aggregation rule satisfies *universal domain*.

The aggregation rule F_{Maj} which accepts only the issues accepted by a (strict) majority of agents is the *majority rule*, and is defined such that for any proposition p, $F_{Maj}(\boldsymbol{B})(p) = 1$ if and only if $\left|N_p^{\boldsymbol{B}}\right| > \frac{n}{2}$. In fact, some aggregation rules will not always output a single ballot, the definition we use is for the *resolute* majority rule which only accepts an issue if there is a strict majority of agents who accept the issue. If we use the non-resolute version of the majority rule, there are cases where the outcome of the rule might be multiple ballots. For example, if $\boldsymbol{B} = (B_1, B_2)$, where $B_1 = (000)$ and $B_2 = (111)$, there is a (weak) majority for both three rejections and three acceptances. Thus $F_{Maj}(\boldsymbol{B}) = \{(000), (111)\}$. To avoid such ties in the outcome, we use the resolute version of the majority rule.[1]

Aggregation Rule Axioms

We can impose conditions on how an aggregation rule should behave. These allow us to compare aggregation rules in a meaningful way and justify the choice of one rule over another. Importantly, these conditions can also help us prove results which hold for larger classes of rules – defined by the conditions they satisfy – rather than for single rules, such as the majority rule. The following conditions or *axioms* for judgment aggregation rules will be relevant for us going forward. An aggregation rule or procedure F satisfies:

- **Collective Rationality** if and only if for any integrity constraint IC, and any profile $\boldsymbol{B} \in \text{Mod(IC)}^{\mathcal{N}}$, $F(\boldsymbol{B}) \in \text{Mod(IC)}$.
- **Consensus Preservation** if and only if $\forall \boldsymbol{B} \in \mathcal{D}^{\mathcal{N}}$, $\forall B^* \in \mathcal{D} : [\forall i \in \mathcal{N} : B_i = B^*] \Rightarrow F(\boldsymbol{B}) = B^*$.
- **Independence** if and only if $\forall p \in \mathcal{I} : \forall \boldsymbol{B}, \boldsymbol{B'} \in \mathcal{D}^{\mathcal{N}} : [\forall i \in \mathcal{N} : B_i(p) = B'_i(p)] \Rightarrow F(\boldsymbol{B})(p) = F(\boldsymbol{B'})(p)$.

Collective rationality says that if all agents submit rational ballots, the outcome of an aggregation rule is always a rational ballot. Consensus preservation states that if every agent submits the same ballot, then the outcome of an aggregation rule should be that same ballot. Independence says that whether an issue is accepted or rejected in the outcome depends only on agents' judgments on that particular issue. If an aggregation rule ignores all but one agent and outputs this agent's ballot as the collective opinion, we call the rule a *dictatorship*. More formally, an aggregation rule F is a dictatorship if $\exists i \in \mathcal{N}$ s.t. $\forall \boldsymbol{B} \in \mathcal{D}^{\mathcal{N}} : F(\boldsymbol{B}) = B_i$.

The complexity of the set of issues is also relevant to whether certain axioms are satisfied in concrete instances. For example, if we know that the agents are only expressing their opinion on a single issue, Independence will

[1] Another way to circumvent this problem is using a tie-breaking rule. For example, we might use lexicographic tie-breaking which in this case would choose the ballot (000) over (111).

not be violated as there is no interplay between multiple issues. For a sufficiently complex set of issues,[2] Dietrich and List [2007a] introduce a generalisation of Arrow's impossibility theorem [Arrow, 1950] in judgment aggregation, which states that any aggregation function with a universal domain which satisfies Collective Rationality, Consensus Preservation, and Independence is a dictatorship.

1.2 Propositional Opinion Diffusion

Propositional Opinion Diffusion (POD) [Grandi et al., 2015] is an attempt to combine ideas about opinion transformations and social influence on a network. POD is an iterative process, where in each step agents in a social network change their opinions based on the opinions of their influencers. The main goal is to characterize when such an iterated diffusion process will terminate, meaning when agents in a social network would stop being influenced to update their opinions by others in the network, and reach a stable opinion.

Notation and Terminology

For a set of agents $\mathcal{N}$, let $E \subseteq \mathcal{N} \times \mathcal{N}$. Then $G = (\mathcal{N}, E)$ is a directed graph where $(i, j) \in E$ means that agent i influences agent j. We call the set of all influence networks $\mathcal{G}$. For a network $G = (\mathcal{N}, E) \in \mathcal{G}$, $\mathit{Inf}(i)_G = \{j \in \mathcal{N} \mid (j, i) \in E\}$ is the set of influencers of agent i on network G. When the network is clear from the context, we will simply write $\mathit{Inf}(i)$. An agent i such that $\mathit{Inf}(i)_G = \emptyset$ is called a *source* in the network G.

Propositional Opinion Diffusion is a discrete time iterative process. At time $t \in \mathbb{N}$ each agent i updates her opinion by aggregating the opinions of her influencers at time $t-1$, using an aggregation rule F_i, which takes as input a profile and returns a ballot, which is agent i's updated opinion at time t. Intuitively this rule tells us in which manner each agent performs her opinion updates. POD uses the framework of binary aggregation to model agents' opinions. We denote with B_i^t the opinion of agent i at time t over a set of issues $\mathcal{I}$, and let $\boldsymbol{B}^t = (B_1^t, \dots, B_n^t)$ stand for the associated profile at time t. We define the iterated process of POD as follows:

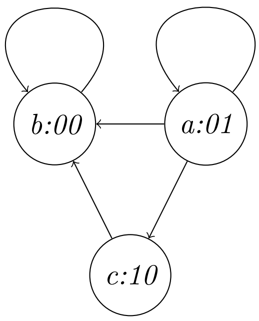
$$B_i^t = \begin{cases} B_i^{t-1} & \text{if } \mathit{Inf}(i) = \emptyset \\ F_i(\boldsymbol{B}^{t-1}_{\mathit{Inf}(i)}) & \text{otherwise} \end{cases}$$

Where $\boldsymbol{B}^{t-1}_{\mathit{Inf}(i)}$ is the profile $\boldsymbol{B}^{t-1}$ restricted to the set $\mathit{Inf}(i)$ of influencers of agent i, and F_i is the aggregation rule of agent i. If $F_i = F$ for all agents $i \in \mathcal{N}$, meaning all agents use the same aggregation rule, the process is called

[2]See definitions of even-number negatability [Dietrich and List, 2007a] and total-blockedness [Nehring and Puppe, 2002].

uniform-POD. The following is a simple example of POD on a social influence network with three agents.

Example 1.1. *Suppose we have three members of a family – agents a, b and c who need to decide on a pet for their household. Their possible options are to either get only a cat (p) or to get only a dog (q) (or choose neither). In other words, their set of possible actions are constrained by the formula* $IC = \neg(p \wedge q)$. *So for any* B_i *(where i is a family member),* $B_i \in Mod(IC) = \{01, 10, 00\}$. *Consider what will happen if these agents are connected in the social influence network below.*



Here agent a is influenced by only herself, agent b is influenced by agents a, c and herself, while agent c is influenced only by a. Their initial opinions are 01 *(getting only a dog),* 00 *(getting neither) and* 10 *(getting only a cat) (of a, b and c respectively), meaning* $\boldsymbol{B}^1 = (01, 00, 10)$. *We assume that each agent uses the majority rule to update her opinions.*

Since agent a is only influenced by herself, her opinion will be stable throughout the diffusion process. Agent b will also not update her opinion in the first iteration as agents a and c disagree on both propositions. Agent c however, will update her opinion by copying the opinion of agent a, her only influencer. Thus $B_c^2 = B_a^1 = 01$. *At the second iteration, both agent a and c will have stable opinions, and agent b will now update her opinion since there is now a majority of agreeing influencers. Thus* $B_b^3 = 01$ *and POD will terminate on a unanimous profile* $\boldsymbol{B}^3 = (01, 01, 01)$, *meaning all three family members agree that they should get only the dog.*

Majority-POD for the (resolute) majority rule is defined formally in the following manner:

Definition 1.2 (Majority-POD)**.**

$$B_i^t = \begin{cases} B_i^{t-1} & \text{if } Inf(i) = \emptyset \\ F_{Maj}(\boldsymbol{B}_{Inf(i)}^{t-1}) & \text{otherwise} \end{cases}$$

Thus, for any agent i, if she has no influencers, she will not change her opinion. This will be the case if the agent is a source in the network. However, if the agent has at least one influencer, the agent will change her opinion to match the majority opinion of her influencers.

1.3 Transformation of Opinion Profiles

Central to the idea of opinion diffusion is the manner in which the change in agents' opinions occurs. List [2011] introduces the concept of a *judgment transformation function* as a way of modeling the change in agents' opinions or judgments. A more formal definition of such functions is given by Grossi and Pigozzi [2014]. A transformation function takes as input a rational profile of ballots and outputs a profile of ballots (which are not necessarily rational). We can view the input profile as the pre-communication or pre-deliberation profile, and the output as the opinions after agents are allowed to deliberate or influence each others' opinions. One reason to consider such transformation functions is that they might be one way of achieving *cohesion-generation*, meaning the application of a transformation function might turn a profile which could not be aggregated using the majority rule into one where the majority rule produces a collectively rational outcome.

Transformation Functions

Let $\mathcal{J} = \langle \mathcal{I}, \mathcal{N}, IC \rangle$ be a binary aggregation problem, A judgment tranformation function for $\mathcal{J}$ is a function

$$T : \mathcal{D}^{\mathcal{N}} \to \mathcal{D}^{\mathcal{N}}.$$

For an individual i and profile $\boldsymbol{B}$, $T_i(\boldsymbol{B})$ denotes the ith set in the transformed profile $T(\boldsymbol{B})$, i.e. agent i's opinion in the transformed profile and $T_{i,p}(\boldsymbol{B})$ for $p \in \mathcal{I}$ denotes i's opinion on p in the transformed profile. We will generally assume that the input profile to a transformation function is in $\text{Mod(IC)}^{\mathcal{N}}$, meaning each of the ballots satisfies the integrity constraint. We can equivalently state the transformation function as a tuple $\langle F_i \rangle_{i \in \mathcal{N}}$ of judgment aggregation rules where F_i is the aggregation rule associated with agent i, which takes as input a profile of ballots and outputs a single ballot – the new opinion of agent i.

One example of a transformation function – variants of which we will examine further in later sections – is *deference to majority*:

$$T_i(\boldsymbol{B}) = F_{Maj}(\boldsymbol{B}).$$

This is the transformation function where in the output, each agent accepts those issues which a majority of agents accepted in the input profile (or the pre-communication profile), meaning for any agent i, $F_i = F_{Maj}$. A transformation function T is an *opinion leader function* if it assigns an opinion leader to each agent, whose pre-communication opinion will become the post-communication opinion of the agent. Thus for any profile $\boldsymbol{B}$ and any agent i, there is some agent j such that $T_i(\boldsymbol{B}) = B_j$. This is equivalent to saying that for every i, F_i is a dictatorship. A transformation function satisfies universal domain if it allows for any rational profile as input.

Conditions For Transformation Functions

For a set of issues $\mathcal{I}$, a set of agents $\mathcal{N}$, and an integrity constraint IC, let $\mathcal{J} = \langle \mathcal{I}, \mathcal{N}, IC \rangle$ be a binary aggregation problem. A transformation function T for $\mathcal{J}$ satisfies:

- **Rationality** if and only if $\forall i \in \mathcal{N}, \forall \boldsymbol{B} \in \text{Mod(IC)}^{\mathcal{N}}$: $T_i(\boldsymbol{B}) \models IC$.
- **Consensus Preservation** if and only if $\forall \boldsymbol{B} \in \mathcal{D}^{\mathcal{N}}, \forall B^* \in \mathcal{D} : [\forall i \in \mathcal{N} : B_i = B^*] \Rightarrow [\forall i \in \mathcal{N} : T_i(\boldsymbol{B}) = B^*]$.
- **Minimal Relevance** if and only if $\forall i \in \mathcal{N} : \exists \boldsymbol{B}, \boldsymbol{B'} \in \mathcal{D}^{\mathcal{N}}$ s.t. $\boldsymbol{B} =_{-i} \boldsymbol{B'}$, $B_i \neq B'_i$ and $T_i(\boldsymbol{B}) \neq T_i(\boldsymbol{B'})$.
- **Independence** if and only if $\forall p \in \mathcal{I}, \forall \boldsymbol{B}, \boldsymbol{B'} \in \mathcal{D}^{\mathcal{N}} : [\forall i \in \mathcal{N} : B_i(p) = B'_i(p)] \Rightarrow [\forall i \in \mathcal{N} : T_{i,p}(\boldsymbol{B}) = T_{i,p}(\boldsymbol{B'})]$.

Rationality requires that each ballot in the outcome of a transformation function satisfies the integrity constraint. Consensus Preservation says that if every agent submits the same ballot, then no one should change their opinion in the outcome of the transformation function. Minimal Relevance states that an agent's own opinion should have some impact on her opinion in the outcome of the transformation function. This means that agents will not always completely ignore their own opinions when performing an update. Independence states that only the opinions on a proposition p should have any impact on whether or not it is accepted by an agent in the outcome.

List [2011] showed that for a sufficiently complex set of issues, the only transformation function with universal domain satisfying Rationality, Consensus Preservation, Minimal Relevance and Independence is the identity function. If we relax the requirement for Minimal Relevance – meaning an agent is free to always ignore her own pre-communication opinion – the results are not much more encouraging. Then, instead of each agent being a dictator (so to speak) of her own opinion change, the transformation function may pick any agent whose pre-communication opinion will dictate her post-communication opinion. Thus, dropping only Minimal Relevance still leaves us with an opinion leader function. Both results rely on the generalisation of Arrows theorem by Dietrich and List [2007a] which was mentioned in the previous section.

Since we are interested in looking at Majority-POD from the angle of transformation functions and not only as an iterated diffusion process, it will be useful to explicitly define the transformation function which each iteration of the propositional opinion diffusion process gives rise to. The Maj-POD function will take as input a social influence network G, a profile $\boldsymbol{B}$, and an agent i. Doing this for each agent $i \in \mathcal{N}$ (for the same network G and profile $\boldsymbol{B}$) will give us the transformed profile after one iteration of POD.

Definition 1.3 (Maj-POD Transformation Function)**.**

$$\mathit{Maj\text{-}POD}(G, \boldsymbol{B}, i) = \begin{cases} B_i & \mathit{if}\ \mathit{Inf}(i) = \emptyset \\ F_{\mathit{Maj}}(\boldsymbol{B}_{\mathit{Inf}(i)}) & \mathit{otherwise.} \end{cases}$$

1.4 Conclusion

In this chapter, we've introduced the terminology and notation for Binary Aggregation with Integrity Constraints, which we will use in the following chapters. We've also started drawing the connection between opinion diffusion and transformation functions. While Propositional Opinion Diffusion is concerned with designing a mechanism for the opinion changes of agents and the role of social influence, the literature on transformation functions focuses more on axiomatic analysis of such mechanisms. Nevertheless, POD, is a specific type of opinion profile transformation. It is therefore natural that we use the framework of transformation functions to analyse how it measures up to other transformations, whether it has certain desirable axiomatic properties and whether it falls under the class of transformation functions covered by the impossibility result in [List, 2011]. The following chapter will be dedicated to taking a closer look at Majority-POD as a transformation function.

2 | Opinion Transformations on a Network

Though Majority-POD works well for opinion diffusion on a single issue, there are several problems that arise when we want agents to give their opinions on several interrelated issues. The main problem is related to what we saw in the example at the beginning of Chapter 1. The outcome of the majority rule does not always satisfy the integrity constraint, depending on the opinions being aggregated. This is a problem in judgment aggregation in general – when the outcome of the majority represents a collective opinion – but even more problematic when it represents the updated opinion of a supposedly rational agent. In addition to the Rationality Axiom, we want to explore other properties which correspond to the axioms we've seen for transformation functions in Chapter 1, but are adapted to transformation of opinion profiles on a network. This will enable us to further explore Majority-POD from an axiomatic point of view.

2.1 POD Transformation Function

We would like to check which of the axioms for transformation functions in Section 1.3 are satisfied by the Majority-POD transformation function. Since we know that the majority rule does not guarantee an outcome which satisfies the integrity constraint and thus fails Collective Rationality, we already know that Maj-POD does not fall under the class of functions covered by the impossibility result by List [2011]. However, since each axiom is still independently desirable, it would also be welcome news that any transformation function we use does not violate too many of them. Since we are transforming opinions on a network, our transformation functions will need to have the influence network as an input in addition to the opinion profile. Thus, we define a transformation function T on a network which takes as input a profile and a network comprised of the same agents which appear in the profile.

$$T : \mathcal{D}^{\mathcal{N}} \times 2^{(\mathcal{N} \times \mathcal{N})} \to \mathcal{D}^{\mathcal{N}}.$$

We first introduce a new axiom that is of relevance to social influence on a network and captures some of the essence of how we want the POD mechanism to work. *Exclusiveness* states that an agent's change in opinion depends only on the opinions of her influencers. This is so we can be sure that an agent is only looking to her influencers before updating her opinion, and does not care about the opinions in the rest of the network. A transformation function T satisfies Exclusiveness if and only if:

- $\forall G, \forall i \in \mathcal{N}, \forall \boldsymbol{B}, \boldsymbol{B'} \in \mathcal{D}^{\mathcal{N}} : [\forall j \in \mathit{Inf}(i) : B_j = B'_j] \Rightarrow [T_i(\boldsymbol{B}, G) = T_i(\boldsymbol{B'}, G)]$.

Proposition 2.1. *Let $\mathcal{J} = \langle \mathcal{I}, \mathcal{N}, IC \rangle$ be a binary aggregation problem on a network $G = (\mathcal{N}, E)$, and T a transformation function. Then the following are equivalent:*

(i) T satisifies Exclusiveness.

(ii) $\forall i \in \mathcal{N}$, $\exists F$ s.t. $\forall \boldsymbol{B} : T_i(\boldsymbol{B}, G) = F(\boldsymbol{B}_{\mathit{Inf}(i)})$.

Where $\boldsymbol{B}_{\mathit{Inf}(i)}$ is the profile $\boldsymbol{B}$ restricted to the set $\mathit{Inf}(i)$, and F is an aggregation rule.

Proof. (⇒) Suppose (i). Then for any two profiles $\boldsymbol{B}$, $\boldsymbol{B'}$ where $B_j = B'_j$ for all $j \in \mathit{Inf}(i)$: $T_i(\boldsymbol{B}, G) = T_i(\boldsymbol{B'}, G)$. Since we know that we can state T as a tuple of judgment aggregation functions, this means that T_i is the judgment aggregation function for agent i. Since T_i ignores the opinions of any $j \notin \mathit{Inf}(i)$, we can easily define a restriction of T_i to only the agents in $\mathit{Inf}(j)$, and obtain an aggregation rule F s.t. $F(\boldsymbol{B}_{\mathit{Inf}(j)}) = T_i(\boldsymbol{B}, G) = T_i(\boldsymbol{B'}, G)$.

(⇐) Suppose (i) does not hold. Then it must be the case that for all $j \in \mathit{Inf}(i)$, $B_j = B'_j$, yet $T_i(\boldsymbol{B}, G) \neq T_i(\boldsymbol{B'}, G)$. Since no agent in $\mathit{Inf}(i)$ changes their ballot, any aggregation rule that is only limited to information from the profile restricted to $\mathit{Inf}(i)$ must give the same outcome for both profiles: meaning for any F, $F(\boldsymbol{B}_{\mathit{Inf}(i)}) = F(\boldsymbol{B'}_{\mathit{Inf}(i)})$, and thus $T_i(\boldsymbol{B}, G) \neq F(\boldsymbol{B}_{\mathit{Inf}(i)})$.

□

We can see that Maj-POD must satisfy Exclusiveness, since each agent uses an aggregation function which takes only the opinions of her influencers as input, and thus (ii) is satisfied. Exclusiveness guarantees that only the influencers of an agent can have an effect on how her opinion is changed. This ensures that the social network does in fact play the role we want it to play in the opinion diffusion process. A transformation function on a network which does not satisfy Exclusiveness would, in a sense, be ignoring the network structure and defeating the purpose of clearly defining an influence relation.

Before we can say anything about how Maj-POD fares with the rest of the axioms we've encountered, the definitions for the axioms have to be slightly altered to include the social network as part of the input to the transformation function.

Conditions for Transformation Functions on a Network

For a set of issues $\mathcal{I}$, a set of agents $\mathcal{N}$, and an integrity constraint IC, let $\mathcal{J} = \langle \mathcal{I}, \mathcal{N}, IC \rangle$ be a binary aggregation problem. A transformation function T for $\mathcal{J}$ on a network satisfies:

- **Rationality** if and only if $\forall G \in \mathcal{G}, \forall i \in \mathcal{N}, \forall \boldsymbol{B} \in \mathrm{Mod(IC)}^{\mathcal{N}}$: $T_i(\boldsymbol{B}, G) \models IC$.
- **Consensus Preservation** if and only if $\forall G \in \mathcal{G}$, $\forall B^* \in \mathcal{D} : [\forall i \in \mathcal{N} : B_i = B^*] \Rightarrow [\forall i \in \mathcal{N} : T_i(\boldsymbol{B}, G) = B^*]$.
- **Minimal Relevance** if and only if $\forall G \in \mathcal{G}, \forall i \in \mathcal{N} : \exists \boldsymbol{B}, \boldsymbol{B'} \in \mathcal{D}^{\mathcal{N}}$ such that $\boldsymbol{B} =_{-i} \boldsymbol{B'}, B_i \neq B'_i$ and $T_i(\boldsymbol{B}, G) \neq T_i(\boldsymbol{B'}, G)$.
- **Independence** if and only if $\forall G \in \mathcal{G}, \forall p \in \mathcal{I}, \forall \boldsymbol{B}, \boldsymbol{B'} \in \mathcal{D}^{\mathcal{N}} : [\forall i \in \mathcal{N} : B_i(p) = B'_i(p)] \Rightarrow [\forall i \in \mathcal{N} : T_{i,p}(\boldsymbol{B}, G) = T_{i,p}(\boldsymbol{B'}, G)]$.
- **Influencer-Unanimity** if and only if $\forall G \in \mathcal{G}$, $\forall B^* \in \mathcal{D}$, $\forall i \in \mathcal{N}$: $[\forall j \in \mathit{Inf}(i) : B_j = B^*] \Rightarrow [T_i(\boldsymbol{B}, G) = B^*]$.
- **Influencer-Independence** if and only if $\forall G \in \mathcal{G}, \forall p \in \mathcal{I}, \forall i \in \mathcal{N}, \forall \boldsymbol{B}, \boldsymbol{B'} \in \mathcal{D}^{\mathcal{N}}$: $[B_i = B'_i$ and $\forall j \in \mathit{Inf}(i) : B_j(p) = B'_j(p)] \Rightarrow T_{i,p}(\boldsymbol{B}, G) = T_{i,p}(\boldsymbol{B'}, G)$.

The first four axioms in this list are exactly the ones which appear in Section 1.3, but have been adapted to include a network G in the input. The fifth axiom – Influencer-Unanimity – states that if a profile $\boldsymbol{B}$ restricted to an agent i's influencers is unanimous, then the outcome of F_i should be the ballot submitted by the influencers. This axiom is stronger than Consensus Preservation, as it preserves consensus limited to the set of influencers of i, and does not require global consensus. Influencer-Independence is a similar restriction of Independence to the set of influencers of an agent i. Together with Exclusiveness, it implies Independence, as Exclusiveness guarantees that the only agents who can influence the outcome of F_i are those in $\mathit{Inf}(i)$. Influencer-Independence on a complete directed graph G is equivalent to the classical Independence axiom for transformation functions. This is, of course, because a complete graph implies that for any i, $\mathit{Inf}(i) = \mathcal{N}$.

Proposition 2.2. *If for all agents $i \in \mathcal{N}$, $T_i(\boldsymbol{B}, G) = \textit{Maj-POD}(G, \boldsymbol{B}, i)$, then T satisfies Influencer-Independence and Influencer-Unanimity, but fails Minimal Relevance and Rationality.*[1]

Proof.

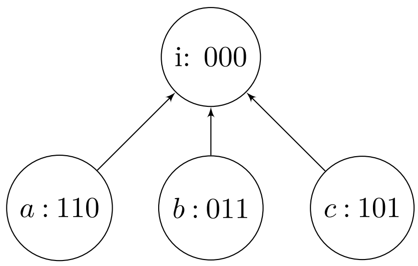
[1] Universal domain holds for any POD transformation function as we allow any profile of rational ballots as input to the function.

- For Influencer-Independence, suppose for some $p \in \mathcal{I}$, $i \in \mathcal{N}$ and two profiles $\boldsymbol{B}$ and $\boldsymbol{B'}$ we have that $\forall j \in \mathit{Inf}(i) : B_j(p) = B'_j(p)$. If $\mathit{Inf}(i) = \emptyset$, then $T_i(\boldsymbol{B}, G) = T_i(\boldsymbol{B'}, G) = B_i$. Otherwise, $T_i(\boldsymbol{B}, G) = F_{Maj}(\boldsymbol{B}_{\mathit{Inf}(i)})$. We know that $\forall j \in \mathit{Inf}(i) : B_j(p) = B'_j(p)$, and that the set of influencers of i are the same in both profiles. It is also known that the majority rule satisfies independence (see [List and Puppe, 2009]), so we know $F_{Maj}(\boldsymbol{B}_{\mathit{Inf}(i)})(p) = F_{Maj}(\boldsymbol{B'}_{\mathit{Inf}(i)})(p)$. Since $T_{i,p}(\boldsymbol{B}, G) = T_{i,p}(\boldsymbol{B'}, G) = F_{Maj}(\boldsymbol{B}_{\mathit{Inf}(i)})(p)$, we conclude that T satisifies Influencer-Independence.

- For Influencer-Unanimity, suppose for an agent $i \in \mathcal{N}$ that all her influencers cast the same ballot, i.e. $\exists B \in \mathcal{D}$ such that $\forall j \in \mathit{Inf}(i)$: $B_j = B$. Then $T_i(\boldsymbol{B}, G) = F_{Maj}(\boldsymbol{B}_{\mathit{Inf}(i)}) = B$, meaning the unanimously submitted ballot will be the opinion of agent i in the updated profile $T(\boldsymbol{B})$.

- Minimal Relevance fails on any network G where there is some agent i, s.t. $i \notin \mathit{Inf}(i)$. Since the POD function satisfies Exclusiveness, we know that $\forall \boldsymbol{B}, \boldsymbol{B'} \in \mathcal{D}^{\mathcal{N}}$ such that $\boldsymbol{B} =_{-i} \boldsymbol{B'}$, it must be the case that $T_i(\boldsymbol{B}, G) = T_i(\boldsymbol{B'}, G)$. Thus there can be no two profiles which differ only with regard to agent i's opinion such that Maj-POD produces different outcomes for i's ballot in the output.

- Finally, and maybe most crucially, Maj-POD does not always produce rational opinions for agents in the output of the transformation function, as the majority rule does not satisfy Collective Rationality and thus may give an outcome which does not satisfy the integrity constraint. So, for $|\mathcal{I}| > 1$, Rationality fails. A simple counterexample shows this.



For the network G above, where $\mathcal{I} = \{p, q, r\}$, let IC $= \neg(p \wedge q \wedge r)$, meaning no agent can accept all of the three propositions. Let a, b and c below be the influencers of some agent i. Then $T_i(\boldsymbol{B}, G) = F_{Maj}(\boldsymbol{B}_{\mathit{Inf}(i)}) = (111) \not\models$ IC. Thus Rationality is violated.

□

Corollary 2.3. *The Majority-POD transformation function satisfies Independence and Consensus Preservation.*

Proof. The corollary is immediate from the fact that Majority-POD satisfies Exclusiveness, Independence, and Influencer-Unanimity. Since Maj-POD satisfies Exclusiveness, we know that any agent not in *Inf*(i) cannot possible influence the outcome of the transformation in any profile. Since we know Influencer-Independence is satisfied, Independence must be satisfied as well, as none of the opinions of non-influencers have any impact on the outcome of the transformation of agent i's opinion. Consensus Preservation is satisfied because if every agent submits the same ballot B^* then it must also be the case that for any agent i, no matter who her influencers are, that they all have submitted the ballot B^*. Since Influencer-Unanimity is satisfied, B^* will be the ballot of i in the outcome of the transformation, and this will hold for all agents. □

It is a problem that Majority-POD fails to satisfy Rationality because it allows agents to hold opinions which do not satisfy the integrity constraint despite the fact that a fundamental requirement in our framework is that agents in the network are rational and hold rational opinions. Additionally, since POD is an iterative process, agents' irrational opinions will potentially need to be aggregated anew in a future iteration of the diffusion process. This means that not only will Majority-POD have irrational agent opinions in the outcome, these outcomes will become inputs to the transformation function in another step. Though the transformation function is defined for a universal domain, it makes little sense when the input to the function is not rational, as the individual aggregation rules are designed to aggregate rational opinions.

We can show that a similar result to the impossibility result in [List, 2011] holds for opinion transformations on a network, where we do not assume that each agent has access to the full information in the network. We do this by replacing Independence with the two axioms of Exclusiveness and Influencer-Independence and Consensus Preservation with Influencer-Unanimity.

Proposition 2.4. *If a transformation function* $T = \langle F_i \rangle_{i \in \mathcal{N}}$ *satisfies Universal Domain, Rationality, Exclusiveness, Influencer-Independence and Influencer-Unanimity, then for any* $i \in \mathcal{N}$*,* F_i *satisfies Collective Rationality, Independence and Consensus Preservation.*

Proof. Suppose there is some T which satisfies Universal Domain, Rationality, Exclusiveness, Influencer-Independence, and Influencer-Unanimity. Recall that we can equivalently define T as a tuple $\langle F_i \rangle_{i \in \mathcal{N}}$ where each F_i is a function, which takes as input $\boldsymbol{B}$ and outputs a ballot. Since T satisfies Exclusivity, we know from Proposition 2.1 that each F_i is equivalent to an aggregation function which only takes as input the profile restricted to *Inf*(i). Thus we can assume that $T_i(\boldsymbol{B}, G) = F_i(\boldsymbol{B}_{Inf(i)})$ for all agents i. Since T satisfies Universal Domain, Rationality, Influencer-Independence, and Influencer-Unanimity, each F_i must satisfy Universal Domain, Collective Rationality, Independence, and Consensus Preservation (defined for aggregation functions). This is immediate from the definitions of Rationality, Influencer-Independence and Influencer-Unanimity. □

Dietrich and List [2007a] tell us an aggregation rule defined for a universal domain which satisfies Collective Rationality, Independence and Consensus Preservation is a dictatorship of one individual.[2] Thus a transformation function which satisfies the axioms in Proposition 2.4 must be an opinion leader function, as each individual has a dictator whose pre-transformation opinion becomes their post-transformation opinion. Proposition 2.4 emphasizes how the impossibility result by List [2011] depends in large part on the axioms satisfied by the individual aggregation rules used by the agents. Thus, even though it, in a sense, generalizes the context in which the impossibility holds – allowing agents to update their opinion based on local, and not global, information – it relies on the same principle as the original impossibility result where each agents uses the information in the entire network. However, the proposition also tells us that Maj-POD does not fall under the class of functions covered by the impossibility, as it does not satisfy Rationality.

2.2 Domain Restrictions

One line of defence against opinions which do not satisfy the IC is to restrict the domain of the aggregation function. This means restricting which profiles are allowed for the initial opinion (of either all agents or a subset of them). *Domain restrictions* such as single-peaked preferences [Black, 1948] have been well-studied in Preference Aggregation, and have gotten some attention in Judgment Aggregation as well [Dietrich and List, 2010]. The particular domain restriction we will examine here is called *unidimentional alignment* of profiles [List, 2003], and it is closely related to the notion of single-peakedness of preference profiles in Preference Aggregation.

A profile $\boldsymbol{B}$ satisfies unidimentional alignment if there exists a strict linear ordering $\prec$ such that for each $p \in \mathcal{I}$ either $N_p^{\boldsymbol{B}} \prec \overline{N_p^{\boldsymbol{B}}}$ or $\overline{N_p^{\boldsymbol{B}}} \prec N_p^{\boldsymbol{B}}$ (where $N \prec N'$ is shorthand for $\forall i \in N, \forall i' \in N' : i \prec i'$). Informally, this means that there is a way of arranging the agents from left to right (or top to bottom) such that for every proposition $p \in \mathcal{I}$, the agents accepting p are either all to the right or all to the left of those rejecting p. We call an agent m the *median agent* wrt. $\prec$ if the sets $\{i \in \mathcal{N} \mid i \prec m\}$ and $\{i \in \mathcal{N} \mid m \prec i\}$ are of the same size. In our results, we will make the assumption that the number of agents is odd to simplify the proofs.[3] Example 2.5 shows a unidimentionally aligned profile.

[2] As before, this holds for a sufficiently complex set of interrelated issues.

[3] Cases with an even number of agents lead to two median agents in the outcome and differ only in that they require a tie-breaking procedure to choose one of the agents.

Example 2.5.

	p_1	p_2	p_3
Agent 1	*0*	*1*	*1*
Agent 2	*0*	*1*	*0*
Agent 3	*1*	*0*	*0*

Unidimentionally aligned profiles are easier to handle, and lead to some positive results about the outcome of the majority rule.

Proposition 2.6 (List, 2003)**.** *Let F be the majority rule, and let $\boldsymbol{B}$ be a profile with an odd number of agents such that $\boldsymbol{B}$ is unidimentionally aligned. Then $F(\boldsymbol{B}) = B_m$, where m is the median agent.*

Let $\mathcal{D} = \{0,1\}^{\mathcal{I}}$ be the domain associated with a set of issues $\mathcal{I}$. We say $D \subseteq \mathscr{P}(\mathcal{D})$ is a *domain restriction* for profiles. We write $\text{sat}(B,D)$ to mean that a profile $\boldsymbol{B}$ is in the restricted domain D. D is *closed under majority* if $\text{sat}(\boldsymbol{B},D)$ implies $\text{sat}((F_{Maj}(\boldsymbol{B}),\boldsymbol{B}),D)$, where $(F_{Maj}(\boldsymbol{B}),\boldsymbol{B})$ is the profile $\boldsymbol{B}$ with the added ballot $F_{Maj}(\boldsymbol{B})$. D is *closed under subprofiles* if $\text{sat}(\boldsymbol{B},D)$ implies $\text{sat}(\boldsymbol{B'},D)$ for any $\boldsymbol{B'} \subset \boldsymbol{B}$. Where $\boldsymbol{B'} \subset \boldsymbol{B}$ means $\boldsymbol{B'}$ is the profile $\boldsymbol{B}$ with one or more ballots removed. Note that domain restrictions allow for profiles of different sizes. This is key since closure under majority and subprofiles involves both adding and removing ballots while still remaining in the restricted domain.

Proposition 2.7. *Unidimentional alignment is closed under majority and closed under subprofiles.*

Proof. Let $\boldsymbol{B}$ be a unidimentionally aligned profile, meaning it is in the domain of unidimentionally aligned profiles D. We first show that D is closed under majority. By Proposition 2.6, $F_{Maj}(\boldsymbol{B})$ is the ballot of the median agent (or one of the two medians if we don't restrict ourselves to an odd number of agents), meaning it is equal to one of the ballots already in $\boldsymbol{B}$. We call this ballot B_m, and the new profile obtained is $\boldsymbol{B'} = (B_m, \boldsymbol{B})$. We know for any $p \in \mathcal{I}$ either $N_p^{\boldsymbol{B}} \succ \overline{N_p^{\boldsymbol{B}}}$ or $\overline{N_p^{\boldsymbol{B}}} \succ N_p^{\boldsymbol{B}}$. Suppose, without loss of generality, that it is the former, meaning $\forall i \in N_p^{\boldsymbol{B}}, \forall i' \in \overline{N_p^{\boldsymbol{B}}} : i \succ i'$. Suppose (again without loss of generality) that $m \in N_p^{\boldsymbol{B}}$. Since $\forall i \in \overline{N_p^{\boldsymbol{B}}} : m \succ i$ and none of the other ballots change or are removed, it must remain the case that$N_p^{\boldsymbol{B'}} \succ \overline{N_p^{\boldsymbol{B'}}}$ and thus $\boldsymbol{B'}$ is unidimentionally aligned.

Showing closure under subprofiles is a much simpler task. Let $p \in \mathcal{I}$ and suppose, without loss of generality, that $\forall i \in N_p^{\boldsymbol{B}}, \forall i' \in \overline{N_p^{\boldsymbol{B}}} : i \succ i'$. Since this is a universal statement about all agents, it will clearly still hold if any agents are removed, so any subprofile of $\boldsymbol{B}$ is unidimentionally aligned. □

Now we are prepared for the main result of this section. We combine our knowledge about influence networks without cycles – directed acyclic graphs – and unidimentionally aligned profiles, to obtain some positive termination results. We say the opinion of agent i is *stable* at time t if at any time $t' \geq t$, $B_i^{t'} = B_i^t$. We say POD has terminated on a network if the opinion of all agents is stable at time t. Let $S_G = \{j \in \mathcal{N} \mid \mathit{Inf}(j) = \emptyset\}$ be the set of source agents in G. Let $\mathit{diam}(G)$ be the length of the longest path in G.

Theorem 2.8. *For a profile $\boldsymbol{B}$ and a directed acyclic graph G, if $\boldsymbol{B}_{S_G}$ is unidimentionally aligned, then Majority-POD will terminate after at most $\mathit{diam}(G)$ steps, and the profile at termination will satisfy unidimentional alignment.*

Proof. We assume for simplicity that all agents have an odd number of influencers, to avoid dealing with more than one median agent. Let $d(i)$ be the maximal distance from i to a source node. Suppose $\boldsymbol{B}_{S_G}$ satisfies unidimentional alignment, meaning the profile restricted to the source nodes satisfy unidimentional alignment. We assume that the number of influencers of an agent are odd.

- If $d(i) = 0$, then i is a source node, and will not change her opinion as she has no influencers.

- Assume that all nodes j s.t. $d(j) \leq k$ have stabilized at step k of the diffusion process, and that the profile $\boldsymbol{B}^k$ at time k restricted to only those agents j such that $d(j) \leq k$ is unidimentionally aligned. Let i be an agent s.t. $d(i) = k+1$. Since we are on a directed acyclic graph, it must hold that for any $j \in \mathit{Inf}(i) : d(j) \leq k$, so all of agent i's influencers must have stabilized their opinion by step $k+1$. Further, we know that the set of all agents j s.t. $d(j) \leq k$ must satisfy unidimentional alignment, so any subset of them (specifically $\mathit{Inf}(i)$) must also satisfy the condition, since unidimentional alignment is closed under subprofiles. Since (by assumption) i has an odd number of influencers, $F_i(\boldsymbol{B}^{k+1}_{\mathit{Inf}(i)}) = B_m$ where B_m is the ballot of the median agent among $\mathit{Inf}(i)$. After this step, the profile $\boldsymbol{B}^{k+1}$ restricted to agents i s.t. $d(i) \leq k+1$ is unidimentionally aligned, since unidimentional alignment is closed under majority.

For any agent i, $d(i) \leq \mathit{diam}(G)$. Thus, at step $\mathit{diam}(G)$, POD will terminate on a unidimentionally aligned profile. □

Essentially what the theorem states is that on a directed acyclic graph, all we need is for the source agents to agree on a linear ordering over the propositions in order for this to propagate down the network and guarantee termination on a rational profile. This result is encouraging not only because it prohibits irrational ballots, but because we know that the unidimentionally aligned profile we are left with at termination can be aggregated without problems using the majority rule.

In fact the statement of Theorem 2.8 holds not just for unidimentionally aligned profiles, but, as the proof suggests, for any domain restriction which is closed under majority and under subprofiles. While restricting the domain is one way of avoiding agents ending up with opinions which are irrational, agenda restrictions are in some cases not a very realistic assumption as they limit the set of opinions we allow agent to have beyond simply disallowing irrational ones. Additionally, the intuitive appeal of these domain restrictions in judgment aggregation can be challenged, as they are not always as natural to impose in our setting as they are in preference aggregation. However, if we are working with all ballots which satisfy the IC (and not with a restricted domain), we are bound to run into cases where the majority rule will give an outcome which does not satisfy the integrity constraint.

2.3 Conclusion

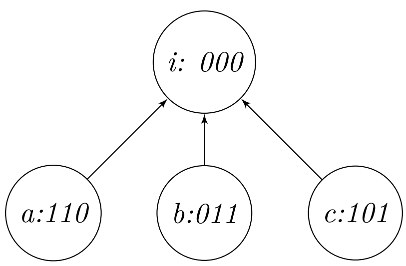
We can see that analysing Majority-POD using an axiomatic approach can give us some insight into how the Propositional Opinion Diffusion process relates to other ways of transforming opinion profiles. In this chapter, we have showed that the introduction of the social influence network does not rid us of the impossibility result by List [2011], as the impossibility depends in large part on the fact that each agent is using an aggregation rule to update her opinion. Crucially, we've seen that Maj-POD does not satisfy the Rationality axiom, which poses a problem to us as it leads to violation of the rationality requirements we impose on agents opinions. We have also see one way to avoid this difficulty, by restricting the domain of the aggregation rules used by agents to update their opinions. Specifically, we've showed that if the sources in a network belong to a domain which is closed under majority and subprofiles, then this property will propagate through the network in directed acyclic graphs and result in a profile which belongs to this same restricted domain.

3 | Propositionwise Opinion Diffusion

As we have seen, the majority rule in judgment aggregation may in some cases produce inconsistent outcomes if we are allowing a universal domain. This is problematic for many reasons even when using majority to reach a rational collective decision, but for our purposes it poses an added problem since we would like the outcome of the rule to be a permissible opinion which can be aggregated in a future iteration of the opinion diffusion process. Thus, we are faced with a problem that was not present when aggregating opinions on a single issue. As we always require agents to be individually rational (i.e. have their opinions satisfy the constraint), we cannot allow them to assume opinions they know are not. Our only other option, then, is to block any change at all if the outcome of the majority rule does not satisfy the constraints. However, this is a quite limiting framework that excludes opinion updates that we might intuitively want to allow.

When working with social influence networks, the structure of the network plays a part in determining whether the majority outcome will be consistent. If for example, we add the assumption that an agent's influencers always have similar opinions which can be aggregated in a consistent way, this is a possible way to circumvent the problem, as we saw was the case with certain domain restrictions. This approach does not allow for networks where we don't necessarily know the agents' opinions ahead of time, however. If we allow for all opinions in the initial profile however, then for some influence networks, the integrity constraint will block all opinion updates if the POD framework will not let an agent assume an opinion that does not meet the integrity constraint. Consider for example the following network,

Example 3.1. *Let* $\mathcal{I} = \{p,q,r\}$, $\mathcal{N} = \{a,b,c,i\}$, *and* IC$= \neg(p \wedge q \wedge r\}$.



As agents a, b and c are sources, they have no influencers and their opinions will not change under POD. However, agent i has an opinion that is exactly the opposite of the majority opinion of her influencers *on all issues*. It makes sense, then, that there should be some change to i's opinion if she is looking to her influencers with the goal of updating her opinion. Clearly she cannot simply assume the majority opinion as it does not satisfy the constraint, which means if we block irrational updates, agent i can never change her opinion and the termination profile will simply be the initial profile. In cases like this, POD over multiple issues is too restrictive in terms of what opinion updates it allows.

Most of us know that when we are influenced to change our opinion, it is rare that we do so on *all* of our opinions. It makes intuitive sense therefore, to allow agents to pick which issues they are willing to be influenced on at any time, or rather, on which issues they will turn to their friends for advice at any step of the opinion diffusion process. This is the main idea behind *Propositionwise Opinion Diffusion*. Suppose for example I am entirely sure about my opinion on some issue p and not willing to change my opinion on this issue at this point in the deliberation. I am however, willing to look to my friends for advice on some other (possibly related) issue q. Crucially, if the aggregate opinion of my friends on q is not compatible with my own opinion on p (meaning changing my opinion on *only* q and not p will violate the integrity constraint), my own opinion on p will be given more weight. In other words, if I either have to change my own belief on an issue other than q or not update my opinion on q at all, I will always choose the latter option.

Propositionwise Opinion Diffusion (PWOD) models precisely this type of opinion change where agents do not change their opinion over the entire set of issues in each iteration. It is an iterative diffusion process where at each step, agents are willing to look to their influencers opinions on only a subset of the issues. In this thesis, we will only consider the case where agents ask advice on a single proposition at each turn, but the framework can easily be extended to accommodate agents seeking advice on any number of issues at any step of the diffusion process. A key difference between POD and PWOD is that PWOD always blocks opinion updates where the outcome of the update does not satisfy the integrity constraint. Thus, the guaranteed satisfaction of the Rationality axiom for transformation functions is built into the diffusion process itself.

Our examination will mainly focus on cases where agents would like to attempt to update their opinion on all propositions, but differ in the order they would like to do this. This allows for a comparison between POD and PWOD, and allows us to tease out the strengths and weaknesses of each.

3.1 Framework & Notation

As before we have a set of issues $\mathcal{I}$, a set of agents $\mathcal{N} = \{1,\dots,n\}$, and an influence network $G = (\mathcal{N},E)$ where $(i,j) \in E$ means agent i influences agent j and $\mathit{Inf}(i)_G = \{j \in \mathcal{N} \mid (j,i) \in E\}$ is the set of influencers of agent i in the network G. We again denote with B_i^t the opinion of agent i at time t, and let $\boldsymbol{B}^t = (B_1^t,\dots,B_n^t)$ stand for the associated profile at time t. We write $\boldsymbol{B} =_{-i} \boldsymbol{B'}$ to mean the profiles $\boldsymbol{B}$ and $\boldsymbol{B'}$ are identical if we ignore agent i's ballot. We write $B =_{-p} B'$ to mean that the ballots B and B' are identical on every issue if we ignore the proposition p, i.e. $\forall q \neq p \in \mathcal{I} : B(q) = B'(q)$. We use the same notation for profiles. Let $\mathit{flip}(B,p)$ be the ballot resulting from changing the judgment on p in the ballot B. For example, if $B(p) = 0$ and $B' = \mathit{flip}(B,p)$, then $B' =_{-p} B$ and $B'(p) = 1$.

We define a PWOD function, which takes as input a network G, a profile of opinions $\boldsymbol{B}$, an agent i, and the proposition p which agent i wants to update. The function returns the updated opinion of i according to F_i – the aggregation rule agent i is using – if such an update continues to satisfy the integrity constraint.

Definition 3.2.

$$PWOD(G,\boldsymbol{B},i,p) = \begin{cases} \mathit{flip}(B_i,p) & \text{if } F_i(\boldsymbol{B}_{\mathit{Inf}(i)})(p) \neq B_i(p) \text{ and } \mathit{flip}(B_i,p) \models IC \\ B_i & \text{otherwise.} \end{cases}$$

The main difference between this type of opinion update and the one we've see in POD is that the satisfaction of the integrity constraint is built into the update function. Thus, we are guaranteed that no agent will hold an irrational opinion at any point of the diffusion process.

A natural restriction to define is the one where every agent uses the majority rule to aggregate the opinion of her influencers on a given issue p. We call this function Majority-PWOD, or Maj-PWOD, and it will be the main point of exploration in this thesis. The Maj-PWOD function is the PWOD function we saw above, but with $F_i = F_{\text{Maj}}$ for all agents $i \in \mathcal{N}$.

Definition 3.3.

$$\mathit{Maj\text{-}PWOD}(G,\boldsymbol{B},i,p) =$$

$$\begin{cases} \mathit{flip}(B_i,p) & \text{if } F_{\mathit{Maj}}(\boldsymbol{B}_{\mathit{Inf}(i)})(p) \neq B_i(p) \text{ and } \mathit{flip}(B_i,p) \models IC \\ B_i & \text{otherwise.} \end{cases}$$

Having defined how each agent makes changes to her opinion at any step, we can now define when a transformation is permissible for a given aggregation problem, meaning when a change in opinion will not violate the integrity constraint. For a given network G an integrity constraint IC, a set of F_i for $i \in \mathcal{N}$, and any two profiles $\boldsymbol{B}$ and $\boldsymbol{B'}$, we say there is a *permissible transformation* from $\boldsymbol{B}$ to $\boldsymbol{B'}$ (on G) if: $\exists S \subseteq \mathcal{N}$ (where $S \neq \emptyset$) where $\forall i \in S : \exists p \in \mathcal{I}$ such that $\text{PWOD}(G, \boldsymbol{B}, i, p) = B'_i$. We say the transformation is *effective* if there is some $i \in S$ such that $B_i \neq B'_i$.

Informally this means that there is an (effective) permissible transformation from the profile $\boldsymbol{B}$ to the profile $\boldsymbol{B'}$ if we can find a nonempty subset of agents who each change their opinion on one proposition in the profile $\boldsymbol{B}$, and this results in reaching the profile $\boldsymbol{B'}$. The notion of a permissible transformation is defined *for* a given integrity constraint and after fixing F_i for each agent i.

3.2 Termination of PWOD

For a given network G and an agent $i \in \mathcal{N}$, we say agent i's ballot B_i^t is *stable* at time t if for all $t' \geq t$, $\forall p \in \mathcal{I}$: $\text{PWOD}(G, \boldsymbol{B^{t'}}, i, p) = B_i^t$. If the only permissible transformation from a profile $\boldsymbol{B}$ is mapping the profile to itself, this means there are no agents who can change their opinion in this profile, and thus that no effective transformation exists. This is equivalent to saying that for every agent $i \in \mathcal{N}$, i's ballot in $\boldsymbol{B}$ is stable. If, during the diffusion process, we reach such a profile, the opinion diffusion will have terminated, as there are no more possible changes for any agent. We define termination as a property of profiles on a given network. We call a profile $\boldsymbol{B}$ a *termination profile* if there is no $\boldsymbol{B'} \neq \boldsymbol{B}$ such that there is a permissible transformation from $\boldsymbol{B}$ to $\boldsymbol{B'}$ (for some given network G) or equivalently, if all ballots in $\boldsymbol{B}$ are stable.

Definition 3.4 (Asymptotic Termination of PWOD)**.** *For a set of agents $\mathcal{N}$ a set of issues $\mathcal{I}$, an integrity constraint IC and an aggregation rule F_i for each $i \in \mathcal{N}$, we say there is asymptotic termination of PWOD on a network G if for any profile $\boldsymbol{B}$, there is a sequence of permissible transformations leading from $\boldsymbol{B}$ to a termination profile.*

Note that Asymptotic termination of PWOD does not guarantee that the process will indeed always terminate on the network G.[1] For example, if the network is a cycle, there is certainly a choice of updates which would ensure a termination profile is reached, but we cannot be sure that any order of updates will lead to termination. Thus we have to define the stronger notion of universal termination.

Definition 3.5 (Universal Termination of PWOD)**.** *For a set of agents $\mathcal{N}$ a set of issues $\mathcal{I}$, an integrity constraint IC and an aggregation rule F_i for each $i \in \mathcal{N}$, we say there is universal termination of PWOD on a network G, if for*

[1] Although if every permissible transformation has a nonzero probability, termination will eventually occur with probability one.

any profile $\boldsymbol{B}$*, any sequence of effective permissible transformations from* $\boldsymbol{B}$ *leads to a termination profile.*

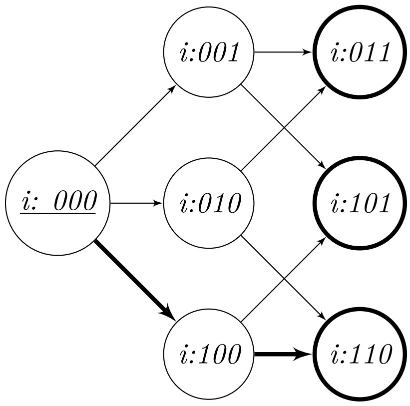
For a set of agents $\mathcal{N}$ a set of issues $\mathcal{I}$, an integrity constraint IC and an aggregation rule F_i for each $i \in \mathcal{N}$, we say PWOD *converges to a unique profile* on a network G if for any two termination profiles $\boldsymbol{B}$ and $\boldsymbol{B'}$, $\boldsymbol{B} = \boldsymbol{B'}$. PWOD *converges to a consensus* on a network G (for set of F_i for $i \in \mathcal{N}$) if for any termination profile $\boldsymbol{B}$, $\forall i, j \in \mathcal{N} : B_i = B_j$.

Note that convergence to a unique profile and convergence to consensus are both defined as properties of termination profiles, meaning they do not require universal termination of PWOD, but rather that, there will be only one termination profile (in the case of convergence to a unique profile), or that all termination profiles will be unanimous (in the case of convergence to consensus). We give an example of (universal) termination which neither satisfies convergence to a unique profile, nor convergence to consensus.

Example 3.6. *Recall Example 3.1 where* $\mathcal{I} = \{p, q, r\}$*,* $\mathcal{N} = \{a, b, c, i\}$*, and IC=* $\neg(p \wedge q \wedge r\}$*. In that case Maj-PWOD will terminate, though it will not converge to a unique profile. Since agents* a, b *and* c *are sources, their opinions are already stable. Agent* i *however has several termination profiles available to her depending on which order she chooses to update her opinions. For example, if she chooses to first update her opinion on* p*, then* q *then* r*, she will first change her opinion to* (100) *(when updating on* p*), then to* (110) *(when updating on* q*). At this stage however, she will be blocked if she tried to update her opinion on* r *to match the majority of her influencers. Thus Maj-PWOD will terminate on a profile where agent* i*'s opinion is* (110)*. However, we can see that if agent* i *chose to update in a different order, say* r, q, p*, then her opinion would stabilise on a different ballot –* (011)*. In either case we know the process will terminate, but depending on the order of updates, the termination profile will differ.*

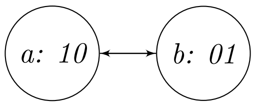
Given this initial profile $\boldsymbol{B}$ *on the network* G*, we can draw a corresponding state transition graph for Maj-PWOD, which allows us to see the effective transformations and the termination profiles. In such a state transition graph, each node represents a different profile on the network* G*, with the underlined node corresponds to the initial profile* $\boldsymbol{B}$ *on this network. An edge (S1, S2) signifies that there is an effective transformation from the state S1 to the state S2.*

Since in this particular case G *is a DAG, there will always be sources who do not change their opinions. Thus, we omit the opinions of agents* a, b *and* c *in our transition graph as they will never change throughout the Maj-PWOD diffusion process. The thicker bordered nodes in the graph represent the termination states for the diffusion problem in the example above. We can also see which states are termination states by checking that they do not have any* *outgoing edges, since a termination state is equivalent to a profile where there are no further effective transformations available. Here we've chosen to use thicker edges to show the transitions which will take place when agent i updates her opinion in a lexicographic order (p,q,r). Note that for a different order of updates, we would reach a different termination state. Note also that this is an example of universal termination as any sequence of effective transformations will lead to a termination profile.*

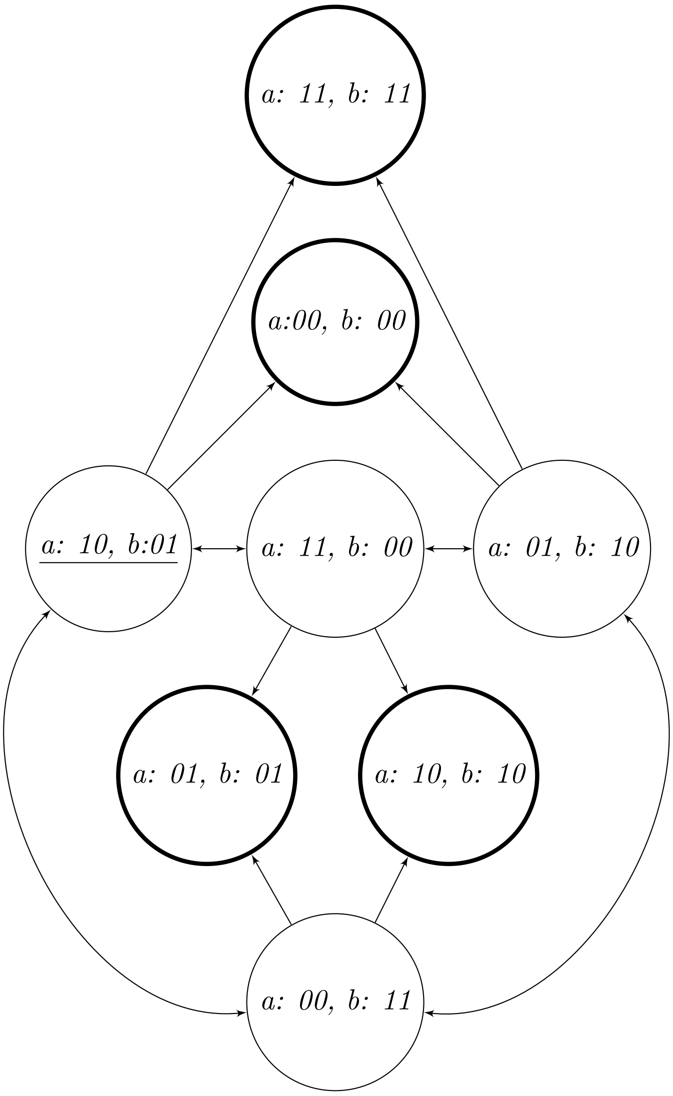


Below is an example of asymptotic termination of Majority-PWOD.

Example 3.7. *Consider the following network G', and initial profile $\boldsymbol{B}$ for agents $\{a,b\}$ and $\mathcal{I} = \{p,q\}$. Let us call this state $S1$.*



The network above is a two-cycle comprising of two agents who are each other's only influencer. This diffusion problem on the social influence graph G' gives rise to the following – more complex – state transition graph. Again the initial profile on G is underlined.



This state transition graph shows us that on the network G, we have asymptotic termination of Maj-PWOD. At any state, there is some sequence of effective transformations which will lead to one of the four termination states. However, we do not have universal termination, as it is possible to perform an effective transformation at each iteration of the process, while never reaching a termination state. Note that for each effective transformation from a state S to another state S' (which is not a termination state), there is also an effective transformation from S' to S. Further, this transition graph shows us that for this particular problem, Maj-PWOD converges to consensus. This is because in all four termination profiles, agents a and b have identical ballots.

Finally, we explicitly define the iterative process of PWOD. We will first need to define two functions. Let *turn*: $\mathbb{N} \to 2^{\mathcal{N}}$ be a *turn function* which tells us which individudals are updating their opinions at any time t. Let $\text{prop}_i : \mathbb{N} \to \mathcal{I}$ be function which tells us which proposition agent i updates at any time t.

On a network G, for a given turn function and a function prop_i for each $i \in \mathcal{N}$:

$$B_i^t = \begin{cases} B_i^{t-1} & \text{if } i \notin \mathit{turn}(t) \\ \text{PWOD}(G, \boldsymbol{B}, i, \text{prop}_i(t)) & \text{otherwise.} \end{cases}$$

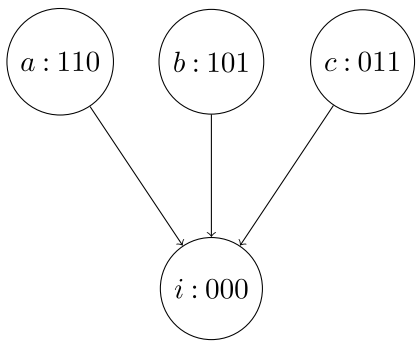
3.3 Conclusion

In this chapter, we presented an alternative to the POD transformation function we saw in the previous chapters, which emphasises the rationality of agents' post-update opinions and allows for a more nuanced way of changing opinions. We defined the formal framework, notation and terminology for Propositionwise Opinion Diffusion, and presented some ways to think about termination of such a process. Mainly, we defined and saw examples of universal and asymptotic termination of PWOD. Finally, we saw that the iterative PWOD process can be defined in a manner similar to POD. We have yet to present any concrete termination results. This will be the focus of the next chapter, where we will explore the role of the integrity constraint in determining what the termination profiles will look like for Majority-PWOD.

4 | Majority PWOD

Part of the reason for wanting to model asynchronous updates on one proposition at a time is to ensure updates will still be possible in cases where the majority of an agent's influencers does not satisfy the integrity constraint. In theory however, we don't know ahead of time whether the IC will be satisfied for every profile just by looking at the network. Since termination profiles are those profiles which tell us the result of the diffusion process if each agent is given the chance to update her opinion on *all* propositions, we would like to know when the proposition-wise process will give the same result as performing the updates on the entire ballot in one step, even when we don't know the particular profile of opinions. More specifically, we would like to ensure that if the outcome of the majority *did* satisfy the integrity constraints, then the propositionwise diffusion process should end up with the same termination profiles as standard POD. We will see that this is not always the case, meaning it is not always possible to update and move towards the majority by changing one proposition at a time, even if the majority satisfies the integrity constraint.

Example 4.1. *The following is an example of an integrity constraint which blocks opinion updates. In this example we have four agents* a, b, c *and* i*. Let* $\mathcal{I} = \{p, q, r\}$ *and* $IC = \neg(p \wedge \neg q \wedge \neg r) \wedge \neg(q \wedge \neg p \wedge \neg r) \wedge \neg(r \wedge \neg p \wedge \neg q)$*, meaning no agent can accept only one of the three propositions.*



Here a, b *and* c *are source nodes and* $Inf(i) = \{a, b, c\}$*. If agent* i *is updating one proposition at a time, then the only opinions, or ballots, she can reach by updating according to the majority opinion of her influencers are* (100)*,* (010) *or* (001)*. All of which violate the IC. Thus each of these updates will be blocked by the integrity constraint and agent* i*'s ballot will never move closer to the majority.*

In Example 4.1 agent i cannot change her opinion at all. In other cases an agent might be allowed to update her opinion towards some majority opinions and not others. Whether updates are going to be allowed or blocked depends in large part on the structure of the integrity constraint.

4.1 Characterization of Integrity Constraints

In this section we present a syntactic characterization of the integrity constraints which do not block opinion updates in the manner we saw in the example above. The integrity constrains we define are precisely those which allow PWOD updates on single issues. This means that if an agents opinion disagrees with the majority opinion of her influencers, she will be able to find one proposition on which they disagree and the integrity constraint will allow her to update her opinion on that proposition to match the opinion of the majority.

We first define the distance between two ballots, and between two formulas. Given two ballots B and $B' \in \mathcal{D}$, the *Hamming distance* between them is $H(B,B') = \Sigma_{p \in \mathcal{I}} |B(p) - B'(p)|$. We can extend this definition to apply to the distance between two formulas which are conjunctions of literals and use the same propositional variables. To distinguish between when we are discussing the distance between these types of formulas (conjunctions of literals) and ballots, we will denote the distance between two formulas ϕ and ψ as $D(\phi,\psi)$. For two conjunctions ϕ and ψ which use exactly the same propositional symbols, we say $D(\phi,\psi) = H(B_\phi, B_\psi)$, where B_ϕ is the unique model of the formula ϕ, restricted to the propositions which occur in ϕ (and ψ). For example, if $\phi = p \wedge q$ and $\psi = p \wedge \neg q$, then B_ϕ is the model in which p and q hold, and B_ψ is the model in which p holds but q does not. In this case $D(\phi,\psi) = 1$.

In order to give a syntactic characterization of the integrity constraints which do not block opinion updates, we transform the formula IC into *full* disjunctive normal form. A formula ϕ is in *disjunctive normal form* if it is a disjunction of conjunctions of literals. In our case a literal is any proposition $p \in \mathcal{I}$, or its negation. A formula is in full disjunctive normal form if each clause uses exactly the same literals [Hein, 2003]. For example, the formula $(p \wedge q) \vee (q \wedge r)$ is in disjunctive normal form, but *not* in full disjunctive normal form, whereas the formula $(p \wedge q) \vee (\neg p \wedge \neg q)$ is in full disjunctive normal form, as the same propositions (or their negations) appear in all clauses of the formula. Any formula of propositional logic can be transformed into (full) disjunctive normal form and for any formula, there might be several equivalent full disjunctive normal form formulas. For example, let $\phi = p \wedge (q \vee \neg q)$. Then, both p and $(p \wedge q) \vee (p \wedge \neg q)$ are full disjunctive normal forms of ϕ.

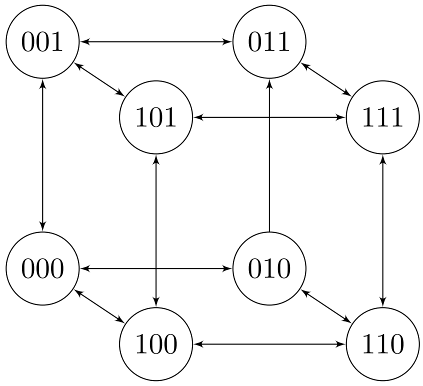
Definition 4.2. *Let IC be an integrity constraint, and* $IC_{DNF} = \phi_1 \vee \cdots \vee \phi_m$

a full disjunctive normal form version of IC. We say IC has an open structure if

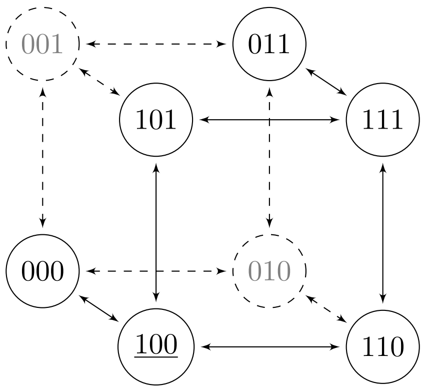
- *for any two disjuncts ϕ, ϕ' of IC there is some sequence $\phi_1, ..., \phi_k$ of (distinct) disjuncts of IC_{DNF} such that $\phi_1 = \phi$, $\phi_k = \phi'$, and for $0 < i < k$, $D(\phi_i, \phi_{i+1}) = 1$ and*
- *$D(\phi_1, \phi_k) = k - 1$.*

Note that each disjunct does not necessarily determine the acceptance or rejection of every issue in $\mathcal{I}$. This is because there is no requirement for the IC to include every proposition in the set of issues. However, we can expand an integrity constraint which does not use all propositions to one which does. For example, if $\text{IC}_{\text{DNF}} = (p \wedge q)$ and $\mathcal{I} = \{p, q, r\}$, we can equivalently state that $\text{IC}_{\text{DNF}} = (p \wedge q \wedge r) \vee (p \wedge q \wedge \neg r)$. Given a set of propositions $\mathcal{I} = \{p_1, \ldots, p_m\}$, and a formula IC in full DNF which uses $k < m$ of these propositions, we can always expand it to a formula which uses all m of the propositions. We say these formulas are in *complete* full DNF. If it is the case that for any two conjuncts ϕ, ϕ' of IC there is a full DNF of IC where is always some path of (distinct) disjuncts $\phi_1, ..., \phi_k$ such that $\phi_1 = \phi$, $\phi_k = \phi'$, and for $0 < i < k$, $D(\phi_i, \phi_{i+1}) = 1$, then this will also hold for the formula in complete full DNF which uses all propositions in $\mathcal{I}$. The key reason for this is that given a number of propositions $k < m$, there is only one full DNF version which uses exactly k of the propositions in $\mathcal{I}$. Thus an equivalent DNF formula with $k+1$ propositions simply has more conjuncts, none of which will increase the distance between disjuncts to more than one. An example makes this clearer. Let IC $\phi \vee \psi$ and suppose IC has an open structure. Expanding it to include another proposition p we get IC $= (\phi \wedge p) \vee (\phi \wedge \neg p) \vee (\psi \wedge p) \vee (\psi \wedge \neg p)$. Clearly if $D(\phi, \psi) = 1$, then $D((\phi \wedge p), (\phi \wedge \neg p)) = 1$ and $D((\phi \wedge p), (\psi \wedge p)) = 1$. So it will still be the case that a path exists between any two disjuncts. We will make use of the complete full DNF of formulas in this chapter. Note though that this is simply to avoid overcomplicating our proofs, and is not a necessary requirement for the results to hold.

The idea behind having an integrity constraint with an open structure, is that it allows for movement between the ballots which satisfy the constraint, when we are using Majority-PWOD. For an integrity constraint with three propositions, p, q and r, we can visualise this connectivity of the models of the IC as a cube. Here, the nodes of the cube represent models, and the edges between them tell us whether it is possible to move from one model to the other by changing the judgment on just one proposition. The cube below shows full mobility between all the nodes, so the integrity constraint is simply $\top$, meaning all ballots are allowed by the constraint.



Note that if we cut some of the edges between the nodes in the cube, we cannot always move freely between the models. This becomes clear in the following example. The dashed edges show which movements are blocked by the integrity constraint, and the dashed nodes show which ballots are not permitted.



We can deduce the integrity constraint from looking at the cube, in this case, IC $= \neg(\neg p \wedge \neg q \wedge r) \wedge \neg(\neg p \wedge q \wedge \neg r)$, meaning q or r cannot be accepted alone. Note that we do not only have to find a path between any two nodes, but a path such that the length of the path is not greater than the distance between the start node and end node plus one. For example, if we start at (000) and our goal is (011), the distance between (100) (underlined) and the end goal of (011), is larger than the distance between (000) and (011). $D(000, 011) = 2$, but $D(100, 011) = 3$. Thus any path from (000) to (011) which includes the node (100) will be longer than $D(000, 011) + 1$. Therefore, we can deduce from this cube that IC does not have an open structure.

In this case PWOD would not allow for the node (100) to be used, even as as a temporary resting point, as we are always limited to changes that reduce the distance between the current ballot and the end goal. Why is this the case? Imagine (011) is the outcome of the majority rule among an agent's

influencers and her own ballot is (000). In order for her to update her opinion using Majority-PWOD, she will have to adopt the opinion of her influencers on some proposition. This means she will have to perform a flip to match the majority. Changing her opinion from (000) to (100) is simply not possible in this case, as she already agrees with the majority on the first proposition. The change in opinion must therefore be on a propositions on which the agents and the majority of her influencers disagree, but as we can see, the only ballots which qualify are the two blocked by the integrity constraint.
Since we are working with the full disjunctive normal form of the integrity constraint, we know that each disjunct will be a conjunction which uses the same propositional atoms and that each proposition which appears in IC will appear once in each disjunct of $\mathrm{IC}_{\mathrm{DNF}}$. The following result will therefore be helpful going forward.

Lemma 4.3. *For a set of agents $\mathcal{N}$ and a set of issues $\mathcal{I}$, if IC is a conjunction of literals, then IC has an open structure and for any $B, B' \in$ Mod(IC), where $B \neq B'$, there is some $p \in \mathcal{I}$ such that $B(p) \neq B'(p)$ and flip$(B,p) \models$ IC.*

Proof. A conjunction of literals means that the disjunctive normal form of IC has only one disjunct – the conjunction itself. So IC must have an open structure.

We now show the latter part of the lemma. Suppose that B and B' are two ballots s.t. $B \neq B'$ and $B, B' \in$ Mod(IC). Since both ballots are models of the IC, they must both accept all propositions that appear as conjuncts in IC. Thus, since $B \neq B'$, they must differ in their judgment on at least one proposition $p \in \mathcal{I}$, where p does not appear in IC. This means that there is no constraint on judgments on p and that whether a ballot accepts or rejects p has no bearing on whether the constraint is satisfied. Thus, if $B \models$ Mod(IC), then it must also be the case that *flip*$(B,p) \models$ IC. □

Since we have showed that an IC which is a conjunction of literals is always open structured, each conjunct of an IC in full disjunctive normal form will always have an open structure, since by definition, each disjunct will be a conjunction of literals. In fact, the open structure of an integrity constraint is what determines whether such flips are permissible in general.

Theorem 4.4. *For a set of agents $\mathcal{N}$ and a set of issues $\mathcal{I}$, IC has an open structure if and only if for any $B, B' \in$ Mod(IC), where $B \neq B'$, there is some $p \in \mathcal{I}$ such that $B(p) \neq B'(p)$ and flip$(B,p) \models$ IC.*

Proof. Suppose IC has an open structure and $\mathrm{IC}_{\mathrm{DNF}} = \phi_1 \vee \cdots \vee \phi_m$ is the integrity constraint in (complete) full disjunctive normal form, which uses all propositions in $\mathcal{I}$. Remember we can do this as if IC has an open structure, then it does not matter which DNF version we are using as the property we are interested in will hold for all of them. Let $B, B' \in$ Mod(IC) be two rational ballots. Then it must be the case that they are also each a model of one of the disjuncts of IC_{DNF}.

For the left to right direction, suppose IC has an open structure. If $B, B' \in$ Mod(ϕ), then $B = B'$ since IC_{DNF} is the integrity constraint in complete full disjunctive normal form. So suppose that $B \in$ Mod(ϕ) and $B' \in$ Mod(ψ). Where ϕ and ψ are distinct disjuncts of $IC_{DNF}$and $D(\phi,\psi) = k$. Since IC uses all the propositions in $\mathcal{I}$, there is only one model of each disjunct of IC_{DNF}. Since IC has an open structure, the exists a path of disjuncts $\phi_1, \ldots, \phi_{k+1}$ where $\phi_1 = \phi$ and $\phi_{k+1} = \psi$, and for any i: $D(\phi_{i+1}, \phi_i) = 1$. Since each disjunct uses each of the propositions in $\mathcal{I}$, they each have a corresponding ballot. This implies there is a series of (distinct) ballots $B_1, \ldots, B_{k+1}$ such that for all $i : H(B_{i+1}, B_i) = 1$. Thus, there must be some ballot $B^* \in$ Mod(IC) which models the disjunct ϕ^* of IC_{DNF}, where $D(\phi, \phi^*) = 1$. Since each disjunct only has one model, we can be sure that $H(B, B^*) = 1$. So there must be some proposition $p \in \mathcal{I}$ s.t. $B^* = \mathit{flip}(B, p)$. Since we know that $D(\phi,\psi) = D(\phi^*,\psi) + 1$, we can conclude that $H(B^*, B') = H(B, B') - 1$. This implies that B and B' must differ on their judgment on p, as otherwise we would not have moved "closer" to B' by flipping the judgment on p in B. In other words, we have that $B(p) \neq B'(p)$ and since $B^* = \mathit{flip}(B, p)$, we also have that $\mathit{flip}(B, p) \models$ IC.

For the right to left direction, we suppose IC does not have an open structure, and present a counterexample to show that it will not hold for any $B, B' \in$ Mod(IC), where $B \neq B'$, that there is always a $p \in \mathcal{I}$ such that $B(p) \neq B'(p)$ and $\mathit{flip}(B, p) \models$ IC. We again assume that the IC_{DNF} is in complete full DNF and uses all propositions in $\mathcal{I}$. Since IC does not have an open structure, it must be the case that for any sequence $\phi, \ldots, \phi'$ of length k, there must be at least two consecutive disjuncts in the sequence such that the distance between then is greater than one. Suppose without loss of generality that these are ϕ and ϕ', and that $D(\phi, \phi') = 2$.[1] Note that this means there is no disjunct ψ of IC_{DNF} such that we can construct the sequence (ϕ, ψ, ϕ'), where $D(\phi, \psi) = 1$ and $D(\psi, \phi') = 1$. Since there is a one-to-one correspondence between the disjuncts and the models of the IC, there must be two ballots B and B' such that $B \models \phi$, $B' \models \phi'$ and $H(B, B') = 2$. Since there is no disjunct ψ which falls between ϕ and ϕ', we can conclude that there is no ballot B^* s.t. $H(B, B^*) = 1$ and $H(B^*, B') = 1$. This means there can be no $p \in \mathcal{I}$ such that $\mathit{flip}(B, p) \in$ Mod(IC) and $B(p) \neq B'(p)$, as this would mean $H(B, \mathit{flip}(B, p)) = H(\mathit{flip}(B, p), B') = 1$. □

Corollary 4.5. *IC has an open structure if and only if for any two ballots* $B, B' \in$ *Mod(IC) such that* $H(B, B') = k$*, there exists a sequence of ballots*

[1] This is the minimal possible distance between the two disjuncts, and in other words, the best scenario case. If we suppose $D(\phi, \phi') > 2$, the proof would proceed in a similar way.

$B_1, \ldots, B_{k+1}$ such that $B_1 = B$, $B_{k+1} = B'$, each ballot in the sequence is a model of the constraint, and for all i: $B_{i+1} = \mathit{flip}(B_i, p)$ for some $p \in \mathcal{I}$.

Proof. We prove the left to right direction by induction on the Hamming distance between B and B'. We know by Theorem 4.4 that if IC has an open structure, then for any two ballots $B, B' \in \text{Mod(IC)}$, there is some $p \in \mathcal{I}$ such that $B(p) \neq B'(p)$ and $\mathit{flip}(B,p) \in \text{Mod(IC)}$. We prove that this implies there is a sequence of (distinct) ballots $B_1, \ldots, B_k$ such that $B_1 = B$, $B_k = B'$, where each ballot in the sequence is a model of the constraint, and for all i: $B_{i+1} = \mathit{flip}(B_i, p)$ for some $p \in \mathcal{I}$.

If $H(B,B') = 1$, then clearly there is only one proposition p on which they disagree, meaning there is a sequence B, B' of length two, where $B' = \mathit{flip}(B,p)$.

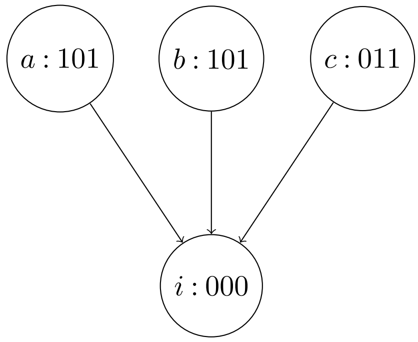
Suppose $H(B,B') = k+1$. By our inductive hypothesis, it holds for any ballot $B^* \in \text{Mod(IC)}$, where $H(B,B^*) = k$ (and $H(B^*,B') = 1$) that there is a sequence of ballots $B_1, \ldots, B_{k+1}$ such that $B_1 = B$, $B_{k+1} = B^*$, where each ballot in the sequence is a model of the constraint, and for all i: $B_{i+1} = \mathit{flip}(B_i, p)$ for some $p \in \mathcal{I}$. Since we know there must be some proposition p s.t. $B(p) = B^*(p) \neq B'(p)$, and $\mathit{flip}(B^*, p) = B'$, this implies that we have a sequence $B_1, \ldots, B_{k+2}$ of length $k+2$, where $B_1 = B$ and $B_{k+2} = B'$ and it hold that for any two consecutive ballots B_i, B_{i+1} that $B_{i+1} = \mathit{flip}(B_i, p)$.

For the right to left direction suppose that for ballots $B, B' \in \text{Mod(IC)}$, we know there is a sequence of ballots $B_1, \ldots, B_{k+1}$, where $B_1 = B$ and $B_{k+1} = B'$ and the Hamming distance between any two consecutive ballots is 1. We simply want to show that there must be some $p \in \mathcal{I}$ such that $B(p) \neq B'(p)$ and $\mathit{flip}(B,p) \models \text{IC}$, since by Theorem 4.4, this implies that IC has an open structure. We know that $B_2 = \mathit{flip}(B,p)$ for some $p \in \mathcal{I}$, so all we have to do is show that $B(p) \neq B'(p)$. Since we know $H(B,B') = k$, $H(B,B_2) = 1$ implies that $H(B_2,B') = H(B,B') - 1$. Of course, the only way this is possible is if $B_2(p) = B'(p)$, or in other words $B(p) \neq B'(p)$. So IC must have an open structure. □

The following is an example of an integrity constraint with open structure, which allows for this type of movement between any ballots which satisfy the given integrity constraint.

Example 4.6. *Let $\mathcal{I} = \{p, q, r\}$ and $IC = (p \wedge q) \rightarrow r$, meaning no agent can accept both propositions p and q and reject r.[2] Then $IC_{DNF} = (p \wedge q \wedge r) \vee (p \wedge \neg q \wedge r) \vee (\neg p \wedge q \wedge r) \vee (p \wedge \neg q \wedge \neg r) \vee (\neg p \wedge q \wedge \neg r) \vee (\neg p \wedge \neg q \wedge r) \vee (\neg p \wedge \neg q \wedge \neg r)$ is the complete full DNF of IC.*

[2]Note that this is the same example we saw in the introductory chapter, with the admissions committee.



In this network, agents a, b and c are sources and Inf$(i) = \{a, b, c\}$*. Note that aggregating the opinions of a, b and c using the majority rule on all issues, will produce an outcome that satisfies the integrity constraint –* (101)*. Further, we can see that at any step of a diffusion process, there will be an issue that agent i can update that will move her closer to the majority of her influencers, and at each step, her opinion will be a model of IC.*

Note that if agent i's opinion is (010)*, she will not be able to update her opinion on p in the first step as this will be blocked by the IC. However once she has updated her opinion on q towards the majority of her influencers, she will be free to accept both p and r.*

At first glance, the condition for an integrity constraint to have an open structure may seem somewhat strict and unnatural. In fact, some examination shows us that it there are several quite reasonable and intuitively justifiable integrity constraints which meet our requirements. We've seen that IC which are conjunctions of literals or disjunctions of literals both have an open structure. For example, the integrity constraint which states that an agent must accept at least one of the issues is exactly the one which is a disjunction of literals. We can imagine a scenario where this would be a useful constraint. Suppose for example a set of agents have co-authored a paper and it has been accepted to a conference. They each have to say which of the authors they would like to attend the conference, but of course, it is required that at least one author attend. Thus, they must accept at least one author who should go to the conference, and they are free to have the opinion that more than one author should attend. Another large class of problems for which the integrity constraint has an open structure are preference aggregation problems.

4.1.1 Preference Aggregation

In preference aggregation, each agent is asked to provide an ordering over a set of alternatives. The rankings provided by the agents are then aggregated into a collective ordering which represents the preferences of the group as a whole. The type of constraint which models such a preference aggregation problem in our binary aggregation framework are perhaps among the most intuitively appealing open structured integrity constraints. To show that this indeed is the case, we first have to provide the translation from preference aggregation to binary aggregation with integrity constraints.

Given a preference aggregation problem defined by a set of alternatives A of alternatives and a set of agents $\mathcal{N}$, we can translate A into a set of binary issues $\mathcal{I}_A = \{p_{ab} \mid (a,b) \in A \times A\}$.[3] A ballot B_i of agent $i \in \mathcal{N}$ corresponds to a preference relation (in our case, a strict linear order) $\succ_i$, s.t. $B_i(p_{ab}) = 1$ if and only if $a \succ_i b$. Each ballot must satisfy the following integrity constaints ($\text{IC}_<$):

- Irreflexivity: $\neg p_{aa}$ for all $a \in A$.
- Transitivity: $p_{ab} \wedge p_{bc} \rightarrow p_{ac}$ for $a,b,c \in A$ (where a,b and c are pairwise distinct).

The integrity constraint $\text{IC}_<$ is the conjunction of the two constraints above. All strict preference orders, or linear orders, satisfy $\text{IC}_<$. Further, we know that if two agents have distinct preference orders over a set of alternatives, then it must be the case that they disagree on at least one pair of alternatives where one of the agents can flip her preference over these two alternatives to agree with the second agent, while her preferences over all other alternatives remain the same. For example, if agent i's preference order is $a > b > c$ and agent i''s preference order is $c > b > a$, then agent i can flip her preference $a > b$ to agree with i', and still prefer both a and b to alternative c. In fact, the following propositions show that will always be possible for any two distinct ilinear orders.

Proposition 4.7. *Let $\succ$ and $\succ'$ be two distinct linear orders on a set of alternatives A. Then $\succ$ and $\succ'$ must disagree on at least one pair of alternatives that are adjacent in $\succ$.*[4]

We rephrase the proposition into the notation of the binary aggregation framework.

Proposition 4.8. *Let B and B' be two distinct ballots over a preference agenda $\mathcal{I}_A$, s.t. both $B \models IC_<$ and $B' \models IC_<$. Then B and B' must disagree on at least one proposition p_{ab} where:*

- *either $B(p_{ab}) = 1$ and there is no $c \in A$ s.t. $B(p_{ac}) = 1$ and $B(p_{bc}) = 0$,*
- *or $B(p_{ab}) = 0$ and there is no $c \in A$ s.t. $B(p_{ac}) = 0$ and $B(p_{bc}) = 1$.*

[3] for each pair (a,b), we only include one of p_{ab} and p_{ba} in the set of issues, as rejecting p_{ab} is equivalent to accepting p_{ba} and vice versa.

[4] This proposition is folklore in the literature, though it has been proved formally by Elkind et al. [2009].

Corollary 4.9. *The integrity constraint $IC_<$ has an open structure.*

Proof. By Theorem 4.4 we know that if for any $B, B' \in \text{Mod}(\text{IC}_<)$, there is some $p \in \mathcal{I}$ such that $B(p) \neq B'(p)$ and $\mathit{flip}(B,p) \models \text{IC}$ then the IC has an open structure. So let B, B' be two ballots in $\text{Mod}(\text{IC}_<)$. Then, by Proposition 4.8, we know they must disagree on some proposition such that either $B(p_{ab}) = 1$ and there is no $c \in A$ s.t. $B(p_{ac}) = 1$ and $B(p_{cb}) = 1$ or $B(p_{ab}) = 0$ and there is no $c \in A$ s.t. $B(p_{ac}) = 0$ and $B(p_{cb}) = 0$. Without loss of generality, suppose it is the former. Let $B^* = \mathit{flip}(B, p_{ab})$. We now show that B^* must satisfy the IC. First, we know irreflexivity is satisified as $B'(p_{ab}) = B^*(p_{ab})$ and $B' \in \text{Mod}(\text{IC}_<)$. To see that transitivity holds, we need to check that $p_{ba} \wedge p_{ac} \rightarrow p_{bc}$ is satisfied by B^*.[5] Since B^* and B agree on all propositions but p_{ab}, we know by Proposition 4.8 that there is no c such that $B^*(p_{ac}) = 1$ and $B^*(p_{bc}) = 0$. So $B^* \models \text{IC}_<$, meaning $\text{IC}_<$ has an open structure. ☐

4.1.2 Convergence to Majority

In addition to the notion of convergence to consensus which we defined in Chapter 3, we define a characteristic of termination profiles which is especially desirable for Maj-PWOD.

Definition 4.10. *We say PWOD will converge to majority on a network G if for a termination profile $\boldsymbol{B}$ it is the case that for any $i \in \mathcal{N}$, $B_i = F_{Maj}(\boldsymbol{B}_{Inf(i)})$.*

This is again a requirement for termination profiles, and does not require universal termination. Convergence to Majority turns out to be a fairly strong requirement, and can only be guaranteed if we know that any majority opinion will satisfy the integrity constraint. Fortunately, we know when this is the case. We say an integrity constraint is *lifted* by a certain aggregation rule if the outcome of the rule will always satisfy the constraint, as long as each individual opinion being aggregated is a model of the integrity constraint. Grandi and Endriss [2011] showed that an integrity constraint IC is *lifted* by the majority rule if and only if IC is a conjunction of 2-clauses, i.e. a conjunction of disjunctions of size two.

The class of integrity constraints for which the majority opinion always satisfies the constraint and the class of integrity constraints which have an open structure are distinct, but have a non-empty intersection. First, we can show that the set of integrity constraints which are lifted by the majority do not necessarily all have an open structure. Consider the following example.

Example 4.11. *Let $IC = (p \vee q) \wedge (\neg p \vee \neg q)$.*

[5] More precicely, since p_{ba} is not formally in the set of issues, we must check that $\neg p_{ab} \wedge p_{ac} \rightarrow p_{bc}$ is satisfied.

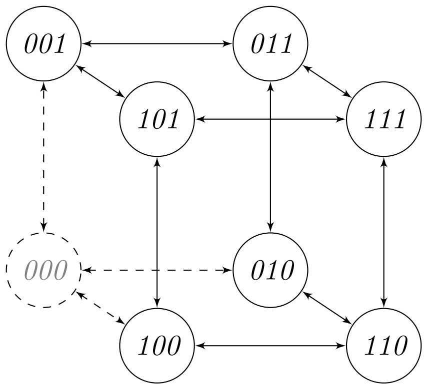
Then $IC_{DNF} = (p \wedge \neg q) \vee (\neg p \wedge q)$. *We can see that the distance between* $(p \wedge \neg q)$ *and* $(\neg p \wedge q)$ *is larger than 1, so an agent could never move from a ballot* $(10) \in$ *Mod*$(p \wedge \neg q)$ *to a ballot* $(01) \in$ *Mod*$(\neg p \wedge q)$ *updating one proposition at a time. Put differently, this integrity constraint blocks both* $(p \wedge q)$ *and* $(\neg p \wedge \neg q)$. *It is also easy to see that IC is a conjunction of 2-clauses, and is therefore lifted by the majority rule.*

So we can say for certain that IC being lifted by the majority rule does not imply that it has an open structure. We now show an example of an integrity constraint which has an open structure but is in fact not lifted by the majority rule to show that the converse does not hold either.

Example 4.12. *Let* $IC = p \vee q \vee r$. *Then clearly IC is not lifted by the majority, as it is conjunction of a single 3-clause.*

However, IC has an open structure. Too see that this is the case, note that we can transform it from DNF (as it is currently a disjunction of conjunctions of size one), to full DNF. This would mean that we expand the disjunct p into $(p \wedge q \wedge r) \vee (p \wedge \neg q \wedge \neg r) \vee (p \wedge \neg q \wedge r) \vee (p \wedge q \wedge \neg r)$ *and similarly with q and r.* [6] *Then for any disjuncts where p holds, there is certainly a path of disjuncts to every other disjunct where p holds, and we can go between the disjuncts where p holds and those where q holds via the ones they have in common –* $(p \wedge q \wedge \neg r)$ *and* $(p \wedge q \wedge r)$.

We can visualise this better on the cube:



Since the only ballot excluded by the integrity constraint is (000), *the remainder of the cube is still fully connected. And, equally importantly, for any two nodes, and any path between them which utilises the node* (000), *there is a path of equal distance which does not need to use the blocked node.*

[6] The end result will be the formula $(p \wedge q \wedge r) \vee (p \wedge \neg q \wedge \neg r) \vee (p \wedge \neg q \wedge r) \vee (p \wedge q \wedge \neg r) \vee (\neg p \wedge q \wedge \neg r) \vee (\neg p \wedge q \wedge r) \vee (p \wedge \neg q \wedge r) \vee (\neg p \wedge \neg q \wedge r)$.

Recall we also said that the intersection of open structured integrity constraints and those lifted by the majority rule was nonempty. A familiar example shows this. If IC is a conjunction of literals then it both has an open structure and is lifted by the majority rule. As we've seen already, an integrity constraint that is a conjunction of literals always has an open structure, and since it is clearly a conjunction of 2-clauses (in fact, each conjunct is simply a propositional atom), it must also be lifted by the majority rule.

4.2 Termination Results for Open Structured IC

If we know that our integrity constraint is open structured, this gives us a lot more information about what a possible termination profile will look like, given certain types of networks.

Definition 4.13 (Chain Graph)**.** *A graph $G = (\mathcal{N}, E)$ is a chain if and only if for any $i, i' \in \mathcal{N} = \{1, \ldots, n\}$: $(i, i') \in E \Leftrightarrow i' = i + 1$.*

Theorem 4.14. *Let G be a chain. Then for any set of issues $\mathcal{I}$ and any set of agents $\mathcal{N}$: if IC has an open structure Maj-PWOD converges to consensus on G.*

Proof. Suppose we have a set of issues $\mathcal{I}$, and a chain graph $G = (\mathcal{N}, E)$, which defines for each $i \in \mathcal{N}$, her set of influencers $\mathit{Inf}(i) = \{i - 1\}$. Let IC be open structured. Suppose for contradiction that Maj-PWOD terminates on a profile $\boldsymbol{B}$ which is not unanimous. Then there must be at least two agents i, i' such that $B_i \neq B_{i'}$ and $i' \in \mathit{Inf}(i)$. Then since IC has an open structure, there must be some $p \in \mathcal{I}$ s.t. $B(p) \neq B'(p)$ and $\mathit{flip}(B_i, p) \models$ IC. Thus, there is a possible transformation, where agent i changes her opinion, and the process has not terminated on $\boldsymbol{B}$. □

A similar result holds for Maj-PWOD on any directed graph G.

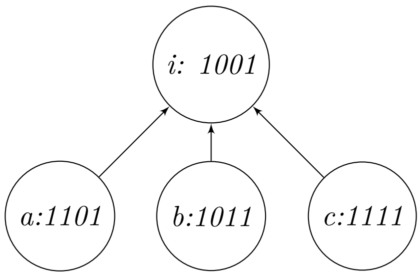
Theorem 4.15. *For an IC which is lifted by the majority rule, Maj-PWOD will converge to majority on any $G = (\mathcal{N}, E)$ if and only if IC has an open structure.*

Proof. Let $G = (\mathcal{N}, E)$ be a social influence network, $\mathcal{N}$ a set of agents, and $\mathcal{I}$ a set of issues. Let IC be such that it is lifted by the majority rule.

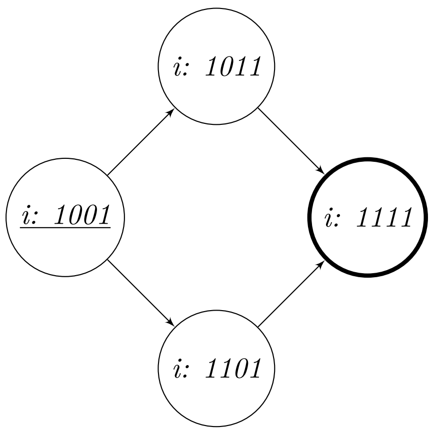
For the left to right direction, suppose IC does not have an open structure. Then, by Theorem 4.4, it must be the case that there are two ballots B, B' such that there is no $p \in \mathcal{I}$ where $B(p) \neq B'(p)$ and $\mathit{flip}(B, p) \models$ IC. Since both B and B' are models of IC, this means $H(B, B') > 1$. Let $B = B_i$ for some agent i in the profile $\boldsymbol{B}$, and $B' = F_{\text{Maj}}(\boldsymbol{B}_{\mathit{Inf}(i)})$. We can always construct such a profile on G, for example by fixing the ballots such that $\forall j \neq i \in \mathcal{N} : B_j = B'$. Then, for agent i, her opinion will be stable on the ballot B_i, as there will be no single proposition which she is able to update, given the opinions of her influencers. Thus, agent i's opinion at termination $B_i \neq B'$, and $\boldsymbol{B}$ is the termination profile. So Maj-PWOD has not converged to majority on G.

For the right to left direction, suppose IC has an open structure. Given some inititial profile $\boldsymbol{B}$, we want to show that Maj-PWOD will terminate on a profile $\boldsymbol{B}^*$ such that for any agent $i \in \mathcal{N}$: $B^*_i = F_{Maj}(\boldsymbol{B^*}_{Inf(i)})$, i.e. that Maj-PWOD will converge to majority on G. Suppose for contradiction that this is not the case, meaning Maj-PWOD will terminate on a profile $\boldsymbol{B'} \neq \boldsymbol{B}^*$. Then there is some agent $i \in \mathcal{N}$ such that $B'_i \neq F_{Maj}(\boldsymbol{B'}_{Inf(i)})$. Since the IC is lifted by the majority, we know $F_{Maj}(\boldsymbol{B'}_{Inf(i)}) \models$ IC. Thus, by Theorem 4.4, there must be some $p \in \mathcal{I}$ s.t. $F_{Maj}(\boldsymbol{B'}_{Inf(i)})(p) \neq B'_i(p)$ and $\mathit{flip}(B'_i, p) \models$ IC. This means there is a possible transformation, where agent i changes her opinion, so $\boldsymbol{B'}$ is not a termination profile. □

Example 4.16. *Let $\mathcal{N} = \{a,b,c,i\}$, $\mathcal{I} = \{p,q,r,s\}$ and $IC = p \wedge s$. Let the following be the initial profile on a social influence network G. Note that this IC is an example of one which is both lifted by the majority rule and has an open structure.*



For Maj-PWOD we obtain the state transition graph below.



Note that there is a single termination state, which corresponds to the profile where agent i's ballot is the outcome of the majority rule on $\boldsymbol{B}_{Inf(i)}$. Thus this is an example of convergence to majority.

In practice, social network which include cycles can cause troubles very easily. Consider for example a network with two agents who perform simultaneous updates and always influence each other on the same issues. We can easily see that what will occur is an infinite swapping of opinions between the two agents. However, on simple cycles, we show that we can still say something about the termination profiles which arise for Majority-PWOD. Note that this result says something about the termination profile itself and not the ease with which the diffusion process will reach it. As we've seen in Example 3.7, there are many permissible transformations on cycles which do not lead to termination profiles.

Definition 4.17 (Simple Cycle)**.** *A graph $G = (\mathcal{N}, E)$ is a simple cycle if and only if for any $i, i' \in \mathcal{N} = \{1, \ldots, n\}$: $(i, i') \in E \Leftrightarrow i' = i + 1$ or ($i = k$ and $i' = 1$).*

Proposition 4.18. *Let G be a simple cycle. Then Maj-PWOD will converge to consensus if IC has an open structure.*

Proof. Suppose for contradiction that Maj-PWOD has terminated on a simple cycle G, on a termination profile $\boldsymbol{B}$ which is not unanimous. Then there must be some $i, j \in \mathcal{N}$ such that $(i, j) \in E$ and $B_i \neq B_j$. Since IC has an open structure, we know there must be some $p \in \mathcal{I}$ such that $B_i(p) \neq B_j(p)$ and $\mathit{flip}(B_j, p) \models IC$. This means the opinion of agent j is not stable. But then there is a permissible transformation from $\boldsymbol{B}$ to $\boldsymbol{B'}$ where $\boldsymbol{B} =_{-j} \boldsymbol{B'}$ and $B'_j = \mathit{flip}(B_j, p)$. This contradicts our initial supposition. So Maj-PWOD must converge to consensus of G. □

4.3 Conclusion

In this chapter, we have provided a syntactic characterization of the class of integrity constraints which allow Propositionwise Opinion Diffusion to proceed without being blocked by the constraint – the constraints which have an open structure. We have shown several termination results, including for chain graphs and cycles. The results in this section are encouraging for aggregation problems where the IC has an open structure. The reasoning behind designing a mechanism like PWOD is to ensure that when the collective majority opinion does not meet rationality requirements, opinion diffusion can still proceed. As we've argued, is is also a more intuitive way of modeling how agents change their opinions based on social influence because it both takes each agent's own opinion into account before they change their opinion, and it puts the rationality of opinions front and center. In addition to these intuitively appealing conditions, our results show that Majority-PWOD retains the strengths of its predecessor Majority-POD when the majority does actually satisfy the constraint. When the constraint is not satisfied by a majority, PWOD avoids the trapfalls of POD, where no opinion update could take place, and gives a justifiable outcome that is as close to the majority as the integrity constraint allows.

5 | Properties of Majority PWOD

One of our starting points in this thesis was the impossibility result by List [2011] regarding the transformation of opinion profiles. Since PWOD is one way of transforming a profile of opinions, we can in fact view Maj-PWOD (and PWOD in general) as a type of transformation function. In order to give a proper analysis of Propositionwise Opinion Diffusion we examine which of the axioms in Chapter 2 are satisfied by the Maj-PWOD transformation function.

Crucially, since the blocking of opinion updates which do not satisfy the integrity constraint is built into the Maj-PWOD transformation function, we know it must satisfy Rationality at the expense of the Independence axiom. Often, Independence is relaxed at the cost of leaving the transformation function open to strategic manipulation by agents. List [2011] argues for relaxing Independence and adopting a more holistic way of transforming opinions, which looks at the agents opinions over the whole set of issues before updating ballots. This way of relaxing independence however, leaves the transformation vulnerable to reasoning about the interplay between issues, which is one reason functions which fail Independence can be manipulated. We will see that although Maj-PWOD fails Independence, the propositionwise aspect of PWOD is a way of avoiding the types of manipulation most commonly excluded by satisfying Independence.

5.1 PWOD Transformation Function

For Maj-PWOD, given a binary aggregation problem $\mathcal{J} = \langle \mathcal{I}, \mathcal{N}, IC \rangle$ and an influence network $G = (\mathcal{N}, E)$, we can define any iteration of the diffusion process as a transformation function. Thus, for a set of agents $S \subseteq \mathcal{N}$ and a proposition $\text{prop}(i) = p \in \mathcal{I}$ which is the proposition agent $i \in \mathcal{N}$ is updating at the current iteration, we can define a transformation function T where:

$$T_i(\boldsymbol{B}, G) = \begin{cases} B_i \text{ if } i \notin S \\ \text{Maj-PWOD}(G, \boldsymbol{B}, i, \text{prop}(i)) \text{ otherwise.} \end{cases}$$

Proposition 5.1. *If a transformation function T is the Maj-PWOD transformation function, then T satisfies Rationality, Consensus Preservation and Minimal Relevance. It also satisfies Influencer-Independence while it fails to satisfy Exclusiveness, Independence and Influencer-Unanimity.*[1]

[1]Again, Universal Domain is always satisfied by any PWOD function as we allow any profile of rational ballots as input.

Proof.

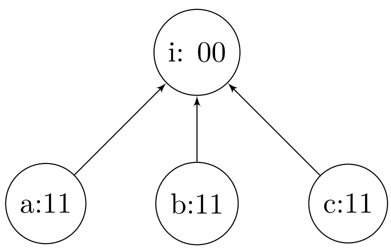
- Rationality is satisfied since if $i \notin S$, then agent i does not make any changes to her ballot, and $T_i(\boldsymbol{B},G) = B_i \models IC$. If $i \in S$, then $T_i(\boldsymbol{B},G) =$ Maj-PWOD$(G,\boldsymbol{B},i,p)$, which by definition of Maj-PWOD, models the integrity constraint, since a *flip* is only performed if the resulting profile satisfies the constraint.

- For Consensus Preservation, suppose $\boldsymbol{B}$ is the profile where every agent submits the same ballot B. Let i be an arbitrary agent, and p the proposition she is updating. If $i \notin S$, then $T_i(\boldsymbol{B},G) = B_i = B$. If $i \in S$, then $T_i(\boldsymbol{B},G) =$ Maj-PWOD$(G,\boldsymbol{B},i,p) = B_i$ since every agent in the network agrees with i on p, and thus, there cannot be a majority among her influencers who reject p and agent i will not change her opinion.

- For Minimal Relevance, we can simply note that agent i must always take into account her own ballot to ensure that changes can be made to her opinion on a subset of $\mathcal{I}$ while still satisfying the constraint.

- Influencer-Independence is satisfied as well. For profile $\boldsymbol{B}$, issue p and an agent i, if p is not the issue she is updating, or if an update to p is blocked by her opinion on the other issues, then $T_{i,p}(\boldsymbol{B},G) = B_i(p)$ and the ballots of all influencers are ignored. So suppose p is the proposition agent i is updating, and $\mathit{flip}(B_i,p) \models$ IC. Let $\boldsymbol{B'}$ be a profile where $B_i = B'_i$ and suppose for all $j \in \mathit{Inf}(i)$ $B_j(p) = B'_j(p)$. Then $F_{Maj}(\boldsymbol{B}_{\mathit{Inf}(i)})(p) = F_{Maj}(\boldsymbol{B'}_{\mathit{Inf}(i)})(p)$, by the definition of the majority rule. Then the definition of Maj-PWOD tells us (since $\mathit{flip}(B_i,p) \models$ IC) if $F_{Maj}(\boldsymbol{B}_{\mathit{Inf}(i)})(p) = B_i(p)$, then $T_{i,p}(\boldsymbol{B},G) = T_{i,p}(\boldsymbol{B'},G) = B_i(p)$. And if $F_{Maj}(\boldsymbol{B}_{\mathit{Inf}(i)})(p) \neq B_i(p)$, then, since $\boldsymbol{B}_{\mathit{Inf}(i)} = \boldsymbol{B'}_{\mathit{Inf}(i)}$, $T_{i,p}(\boldsymbol{B},G) = T_{i,p}(\boldsymbol{B'},G) = F_{Maj}(\boldsymbol{B}_{\mathit{Inf}(i)})(p)$.

- We give a counterexample to show that Exclusiveness, Independence and Influencer-Unanimity fail. Suppose $IC = p \rightarrow q$. Let G be the following network and $\boldsymbol{B}$ the profile shown in the network. Let p be the proposition agent i is updating.



Then $T_i(\boldsymbol{B},G) = B_i$ as an update to p would lead to a ballot which does not satisfy the constraint. Therefore, Influencer-Unanimity fails. Suppose now $B'_i = (11)$ in a profile $\boldsymbol{B'}$, such that $\boldsymbol{B} =_{-i} \boldsymbol{B'}$, then $T_i(\boldsymbol{B'},G) = B'_i$. Thus, Exclusiveness fails, as $T_i(\boldsymbol{B},G) \neq T_i(\boldsymbol{B'},G)$. Suppose now that $B'_i = (01)$ and again $\boldsymbol{B} =_{-i} \boldsymbol{B'}$. Then $T_{i,p}(\boldsymbol{B},G) = 0$, while $T_{i,p}(\boldsymbol{B'},G) = 1$ so Independence fails as well.

□

This is good news in relation to Proposition 2.4 as it means Maj-PWOD will avoid the traps of the impossibility result by List [2011]. The failure of Exclusiveness and Independence is caused by the fact that it is built into the definition of Maj-PWOD that it will block any irrational opinion updates. Thus, it is necessary for an agent to always take into account her own opinion even if she is not in her own set of influencers. This guarantees that Maj-PWOD satisfies Minimal Relevance, which is a very intuitively appealing axiom as it states that an agent should at least take her own opinion into consideration before updating her ballot. When Minimal Relevance fails, this allows for agents who would let an influencer decide what their opinion should be post-transformation, while completely ignoring their own existing opinion. Influencer-Unanimity of course fails as the updates are proposition-wise, and thus there is no way to guarantee an entire ballot will be copied in one iteration of the process. Overall however, Maj-PWOD satisfies most of the important axioms we've discussed, and preserves many of the qualities of the majority rule. The axioms which fails do so because the transformation is defined with the Rationality axiom in mind. In general, the axiomatic analysis of Maj-PWOD paints a picture of an appealing opinion diffusion mechanism which satisfies many desirable conditions for opinion transformations.

5.2 Strategic Agents

One reason Independence is such a desirable condition for judgment aggregation rules is that it is a requirement for strategy-proofness [Dietrich and List, 2007b]. This means that no agent can influence the outcome of the rule in her favor by reporting a ballot which does not represent her truthful judgments. This holds for agents who have what are called *closeness-respecting preferences*. A ballot B is at least as close to the ballot B^* as B' is, if for all issues $p \in \mathcal{I}$ where $B^*(p) = B'(p)$, it is also the case that $B(p) = B^*(p)$. For agent i, we say her preference relation $\succeq_i$ respects closeness to B_i if for any two ballots B, B': if B is at least as close to B_i as B', then $B \succeq_i B'$. A well-known example of closeness respecting preferences are *Hamming preferences*. An agent i with Hamming preferences and the preference relation $\succeq$ will weakly prefer a ballot B to the ballot B' ($B \succeq B'$) if and only if $H(B,B_i) \leq H(B',B_i)$, i.e.

when the Hamming distance between her true opinion B_i and the ballot B is less than or equal to the distance between her true opinion and the ballot B'.

For opinion diffusion processes, we would like to define agents' preferences over opinion profiles. Since the agents are influencing opinion changes in each other, it makes sense for each agent to have preferences over what the opinions of the others in the network are. This allows us to talk about goals such as having as many others in the network with the same opinion as you, or preferences over the ballots of the agents you influence. In particular, we want to say something about agents' goal to spread their opinion in the network. We define *closeness-respecting preferences for opinion profiles* in the following manner. For $S \subseteq \mathcal{N}$ and agent i, we say her preference relation over profiles $\succeq_i'$ respects closeness to B_i if her preference relation over ballots $\succeq_i$ respects closeness to B_i and if for any two profiles $\boldsymbol{B}$, $\boldsymbol{B'}$: if $\boldsymbol{B} =_{-S} \boldsymbol{B'}$ and for all $j \in S$ $B_j \succeq_i B_j'$, then $\boldsymbol{B} \succeq_i' \boldsymbol{B'}$.

If an agent can manipulate the opinion change of just one agent in the network towards a ballot which is closer to her own, she will be able to perform a successful manipulation. There is therefore an intimate connection between strategy-proofness of transformation functions, and strategy-proofness of the aggregation rules used by the agents in the network. Since strategy-proofness of aggregation rules requires an additional axiom, Monotonicity, we need a corresponding Monotonicity axiom for transformation functions. A transformation function T satisfies *Monotonicity* if and only if for two opinion profiles $\boldsymbol{B}$ and $\boldsymbol{B'}$ and for all $p \in \mathcal{I}$: if for $i, j \in \mathcal{N}$, $\boldsymbol{B} =_{-j} \boldsymbol{B'}$, $B_j =_{-p} B_j'$, $B(p) = 0$ and $B'(p) = 1$, then $T_i(\boldsymbol{B})(p) = 1 \Rightarrow T_i(\boldsymbol{B'}(p) = 1$. In other words, for any agent i, if they accepted a proposition p in the outcome of a transformation function T applied to a profile $\boldsymbol{B}$, then added support to this proposition in a profile $\boldsymbol{B'}$ should imply that p is accepted by agent i in the outcome of T.

We also define a variant of the Exclusiveness axiom in Chapter 2, which allows for an agent to influence her own opinion in the outcome of a transformation. We say a transformation function satisfies *Almost-Exclusiveness* if an only if:

- $\forall G \in \mathcal{G}, \forall i \in \mathcal{N}, \forall \boldsymbol{B}, \boldsymbol{B'} \in \mathcal{D}^{\mathcal{N}} : [(\forall j \in \mathit{Inf}(i) : B_j = B_j')$ and $B_i = B_i'] \Rightarrow [T_i(\boldsymbol{B}, G) = T_i(\boldsymbol{B'}, G)]$.

The fact that Almost-Exclusiveness is satisified by Maj-PWOD is clear from the definition. An agent's post transformation opinion is determined by only the opinions of her influencers, and whether her own ballot in the input profile will lead to the update being blocked. Additionally the Maj-PWOD transformation function also satisfies the second axiom needed for strategy-proofness.

Proposition 5.2. *The Maj-PWOD transformation function satisfies Monotonicity.*

Proof. Suppose for a network G and two opinion profiles $\boldsymbol{B}$ and $\boldsymbol{B'}$ that for $i,j \in \mathcal{N}$, $\boldsymbol{B} =_{-j} \boldsymbol{B'}$, $B_j =_{-p} B'_j$, $B_j(p) = 0$ and $B'_j(p) = 1$, and further, that $T_{i,p}(\boldsymbol{B}, G) = 1$.

First suppose $j \neq i$. If $j \notin \mathit{Inf}(i)$ we know $T_{i,p}(\boldsymbol{B'}, G) = T_{i,p}(\boldsymbol{B}, G)$ because Maj-PWOD only takes into account the ballots of agents in $\mathit{Inf}(i)$. So we assume $j \in \mathit{Inf}(i)$. Then if $T_{i,p}(\boldsymbol{B}, G) = 1$, this means either that there was a majority of acceptances for p among agent i's influencers, or that there was a majority of rejections but a change in opinion was blocked by the IC. If the former holds, we know, since agent j is the only one changing her opinion, and only changes her opinion on p, that an additional acceptance for p in $\boldsymbol{B'}$ means $T_{i,p}(\boldsymbol{B'}, G) = 1$. If the latter holds, then agent i's opinion on p will remain the same regardless of her influencers opinions, as her opinions on other issues are the reason for this blocking. As we know that the update was blocked and $T_{i,p}(\boldsymbol{B}, G) = 1$, then it must have been the case that $B_i(p) = 1$ and thus $B'_i(p) = T_{i,p}(\boldsymbol{B'}, G) = 1$.

Now suppose $i = j$. The only way $T_{i,p}(\boldsymbol{B'}, G) = 0$ is if there is a majority of rejections for p among agent i's influencers. If $i \notin \mathit{Inf}(i)$, then the only way she can influence the outcome of T is if the update is blocked. But since by assumption $B_i(p) = 0$ and $T_{i,p}(\boldsymbol{B}, G) = 1$ we know this cannot be the case, as the blocking of an update would mean the outcome of T is the same for both profiles. If $i \in \mathit{Inf}(i)$ and the update is not blocked then, since there was is majority among i's influencers who accept p in $\boldsymbol{B}$, there is a majority of acceptances in $\boldsymbol{B'}$ as well. So the Maj-PWOD transformation function satisfies Monotonicity. □

We now define strategy-proofness for transformation functions. Let $(\boldsymbol{B}_{-i}, B'_i)$ be the profile which is identical to $\boldsymbol{B}$ but with B_i replaced by B'_i. A transformation function T is strategy-proof if there is no network G and no agents $i \in \mathcal{N}$ such that $T((\boldsymbol{B}_{-i}, B'_i), G) \succ_i T(\boldsymbol{B}, G)$, where B_i is agent i's truthful opinion.

Theorem 5.3. *Let G be a network such that $(i,i) \notin E$. Any transformation function which satisfies Monotonicity, Influencer-Independence and Almost-Exclusiveness is strategy-proof for Hamming preferences on G.*

Proof. Let T be a Monotonic, Almost-Exclusive and Influencer-Independent transformation function. Recall that for any network $G = (\mathcal{N}, E)$, it holds that $\forall i : (i,i) \notin E$. Suppose agents all have Hamming preferences.

For an arbitrary agent i, if $j \notin \mathit{Inf}(i)$, we know by Almost-Exclusiveness that she cannot influencer i's opinion at all, so suppose $j \in \mathit{Inf}(i)$. By Influencer-Independence, for any proposition p, agent j can only influence agent i's opinion on p in the outcome by altering her pre-transformation opinion on p. There are then two possible cases:

- $B_j(p) = 0$ and $B'_j(p) = 1$. If $T_{i,p}(\boldsymbol{B}, G) = 1$, then we know by Monotonicity that additional support for p will not change the outcome of T_i. If

$T_{i,p}(\boldsymbol{B},G)=0$ then either the outcome of T_i does not change, in which case j will be indifferent between then two outcomes, or $T_{i,p}((\boldsymbol{B}_{-i},B_i'),G)=1$, which is less preferred by j than the outcome were she to truthfully report her opinion.

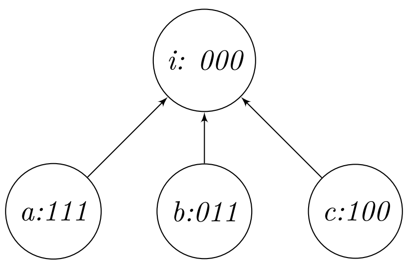
- $B_j(p)=1$ and $B_j'(p)=0$. Again there are two possible outcomes: either no change will occur to the ballot of agent i, and again agent j will be indifferent between the outcome in the two profiles, or $T_{i,p}(\boldsymbol{B},G)=1$ and $T_{i,p}((\boldsymbol{B}_{-i},B_i'),G)=0$, in which case agent j's best strategy would be to report her truthful opinion.

An agent also cannot manipulate her own opinion. The only way i can influence her own opinion is by blocking an update for a single proposition, i.e. by ensuring that $T((\boldsymbol{B}_{-i},B_i'),G)=B_i'$. If $B_i'\neq B_i$, then in the best case for agent i, $H(B_i,B_i')=1$. However, agent i would only change one proposition if her opinion update was not blocked, and thus it is guaranteed that $H(B_i,T_i(\boldsymbol{B},G))=1$, meaning if she has Hamming preferences, it cannot be the case that $T((\boldsymbol{B}_{-i},B_i'),G)\succ_i T(\boldsymbol{B},G)$. □

Corollary 5.4. *The Maj-PWOD transformation function is strategy-proof for Hamming preferences.*

Since Maj-PWOD satisifies Almost-Exclusiveness, Influencer-Independence and Monotonicity, it falls under the class of transformation functions which is strategy-proof for Hamming preferences. However, although each iteration of PWOD is strategy-proof, there are other types of manipulation that it may in fact susceptible to, depending on the level of information available to agents. While each round of opinion updates cannot be strategically manipulated, an agent might have a more long-term manipulation strategy available to her. Consider the following example.

Example 5.5. *Suppose we have a binary aggregation problem with $\mathcal{I}=\{p,q,r\}$ and $IC=p\rightarrow(q\leftrightarrow r)$, meaning if p is accepted, then either both q and r must be accepted or rejected together. Suppose now we have the following social influence network G, with the initial profile $\boldsymbol{B}$.*



Suppose further that we are using Maj-PWOD, and agents have Hamming preferences. Suppose also that agent i updates her opinion on the propositions in a lexicographic order. So the first proposition agent i wants to update is p, then she will go on to update q and finally r. If agent a is aware that this is the order in which agent i is updating her preferences, she will have an incentive to misreport her true opinion.

Consider first what will happen if she announces her true opinion (i.e. accepting all three propositions). Since there will be a majority among i's influencers for p, agent i will change her opinion in the first round of updates to the ballot (100)*. But since the integrity constraint now blocks any changes to q or r (since $IC = p \rightarrow (q \leftrightarrow r)$), the diffusion process will terminate.*

Now consider what will happen if agent a strategically reports the ballot $B'_a = (011)$ on the same network G, let's call this initial profile $\boldsymbol{B'}$. In the first round of updates, there will no longer be a majority for p, meaning agent i will not change her opinion. This leaves her open to accept both q and r in the next two rounds of updates. Thus, the diffusion process will terminate on a profile where agent i's ballot is (011)*. Since agent a has Hamming Preferences over ballots, we know* $(011) \succeq_a (100)$*, and thus, the resulting termination profile will be preferred by agent a over the termination profile which results from her reporting her true ballot.*

Note that agent a's preferences in this example are closeness-respecting. (both for profiles and individual ballots). Despite this, such a manipulation of Maj-PWOD is possible (and, as so much else when it comes to propositionwise opinion diffusion, depends in large part on the integrity constraint). Another necessary condition for this type of manipulation is that the order of updates for agents are determined beforehand and that this information is available to their influencers. If the issue agent i will update at any step of the diffusion process is chosen from the issues in $\mathcal{I}$ uniformly at random, for example, then it will not be possible for agents to manipulate in this way.

5.3 Conclusion

In this chapter, we have given an axiomatic analysis of the Maj-PWOD transformation function, and found that the diffusion mechanism satisfies several desirable properties. Notably, it guarantees the rationality of all ballots in every iteration of the opinion diffusion, which was one of the main problems with Majority-POD. We've also started to explore the effect of strategic agents and showed that while Maj-PWOD does not satisfy Independence – which is commonly needed in judgment aggregation to ensure agents cannot manipulate – each iteration of the diffusion still remains strategy-proof for Hamming preferences. Further, we've seen that the strategy-proofness of each iteration does not extend to the iterative diffusion process as a whole if the order in which an agent updates her opinions is known to her influencers.

6 | Conclusion and Future Work

In this thesis we have examined influence-based opinion diffusion mechanism on social networks using axioms for transformation functions. Central to our examination was the role of the integrity constraint in binary aggregation. The basis for our work was the Propositional Opinion Diffusion (POD) mechanism by Grandi et al. [2015]. In Chapter 2 we provided an axiomatic analysis of the strengths and weaknesses of POD, using the framework of transformation functions which take as input a social network and a profile of opinions and gives the updated opinions on the same network. We found that the Majority-POD transformation function satisfies several important conditions, but has a weakness in that it sometimes will give irrational ballots in the output of the function. In the following chapters we designed a new mechanism, Propositionwise Opinion Diffusion (PWOD) which avoided the main problem of POD – irrational ballots. PWOD enables agents to update opinions on a single issue at a time, and requires that agents take their own opinion into account before being influenced to change their opinion, to ensure that the integrity constraints always remain satisfied. This is key to assuring that no irrational updates will occur. Our main results show that for a small class of integrity constraints – those with an open structure – PWOD terminates on the same profiles as POD when the majority rule outputs only rational opinions on the network. In Chapter 5 we examined the conditions satisfied by the Majority-PWOD transformation function, we found that in addition to Rationality, it satisfies many of the other central axioms we have discussed. Importantly, PWOD satisfies Influencer-Independence, which is the central reason why each iteration of Maj-PWOD is not susceptible to strategic manipulation by agents with Hamming preferences. However, we saw that PWOD is susceptible to a more long term manipulation strategy if agents know the order in which others are updating their opinions. In this respect, POD outperforms PWOD, as is it not susceptible to this type of manipulation.

On one hand, we can interpret our main result in Chapter 4 as positive – they tell us that there is some connection between the termination profiles for POD applied to entire ballots and PWOD applied to individual propositions. On the other hand, since the requirements for an integrity constraint to have an open structure are very strict, this connection is only guaranteed for a small class of constraints. This does give us a point of exploration however, since we know that convergence to majority for all networks and all profiles is not easy to obtain, we might want to ask what types of profiles make this requirement easier to achieve. For example, what happens on networks where agents are only connected to others with similar opinions? In these cases, the majority opinion of an agents influencers would also be closer to her own, and convergence to majority on certain graphs might be possible for a larger class of integrity constraints.

The results in Chapter 4 hold only when all agents want to update their opinion on all propositions in the set of issues. Still, another benefit of a mechanism like PWOD is that the framework easily allows for agents who might be certain about some opinions and only open to influence on a subset of the issues at hand. Suppose for example a group of friends are deciding on a menu for a dinner. It could easily be the case that some are dead-set on their dessert opinions ("tiramisu or nothing!") but willing to be influenced on what they eat for the main course ("I said no to bacalao initially, but my friends are actually making a good point..."). This type of opinion change would result in a different set of termination profiles that the ones we've seen in this thesis. If agents are generally sure about their opinions on different propositions, we might imagine that the opinion diffusion process will terminate on a compromising profile. Further, we have not said much about what a termination profile will look like if the majority does not satisfy the integrity constraint. Although we know that PWOD is able to update opinions on single propositions in some cases where the majority does not satisfy the constraint,[1] In fact, it is still an open problem whether this is possible for Majority-PWOD in all cases, and how close to the majority an agent's opinion can go if there are models of the integrity constraint that are closer to the majority than the current ballot of the agent.

We have also left many types of strategic behaviour unexplored. Grandi et al. [2016] studied the strategic aspects of social influence networks using *games of influence*, where agents can choose to hide or reveal their opinions to others in the network. Additionally, we might want to examine if an agent could manipulate even when there is some randomness involved in which propositions are updated by agents in the network. Allowing these and other types of strategic behavior might lead us to discover what type of strategies agents can use to successfully manipulate the PWOD mechanism.

[1] Recall Example 3.1.

Bibliography

Krzysztof R Apt and Evangelos Markakis. Diffusion in social networks with competing products. In *Algorithmic Game Theory*. Springer, 2011.

Kenneth Arrow. A difficulty in the concept of social welfare. *Journal of Political Economy*, 58:328–346, 1950.

Alexandru Baltag, Zoé Christoff, Rasmus K Rendsvig, and Sonja Smets. Dynamic epistemic logics of diffusion and prediction in social networks. In *Proceedings of the 12th Conference on Logic and the Foundations of Game and Decision Theory*, 2016.

Duncan Black. On the rationale of group decision-making. *The Journal of Political Economy*, 56(1):23–34, 1948.

Markus Brill, Edith Elkind, Ulle Endriss, and Umberto Grandi. Pairwise diffusion of preference rankings in social networks. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence*, 2016.

Zoé Christoff and Jens Ulrik Hansen. A logic for diffusion in social networks. *Journal of Applied Logic*, 13(1):48–77, 2015.

Morris H DeGroot. Reaching a consensus. *Journal of the Americal Statistical Association*, 69(345):118–121, 1974.

Franz Dietrich and Christian List. Arrow's theorem in judgment aggregation. *Social Choice and Welface*, 29(1):19–33, 2007a.

Franz Dietrich and Christian List. Strategy-proof judgment aggregation. *Economics and Philosophy*, 23(03):269–300, 2007b.

Franz Dietrich and Christian List. Majority voting on restricted domains. *Journal of Economic Theory*, 145(2):512–543, 2010.

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

John Dryzek and Christian List. Social choice theory and deliberative democracy: a reconciliation. *British Journal of Political Science*, 2003.

Edith Elkind, Piotr Faliszewski, and Arkadii Slinko. Swap bribery. In *Algorithmic Game Theory*, pages 299–310. Springer, 2009.

Umberto Grandi and Ulle Endriss. Binary aggregation with integrity constraints. In *Proceedings of 22nd International Joint Conference on Artificial Intelligence*, 2011.

Umberto Grandi, Emiliano Lorini, and Laurent Perrussel. Propositional opinion diffusion. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '15, pages 989–997, 2015.

Umberto Grandi, Emiliano Lorini, and Laurent Perrussel. Strategic disclosure of opinions on a social network. In *Proceedings of the 2016 International Conference on Autonomous Agents*, 2016.

Davide Grossi and Gabriella Pigozzi. *Judgment Aggregation: A Primer*. Morgan and Claypool, 2014.

James L Hein. *Discrete Mathematics*. Jones and Bartlett Publishers, 2 edition, 2003.

Jack Knight and James Johnson. Aggregation and deliberation: On the possibility of democratic legitimacy. *Political Theory*, 22(2):277–296, 1994.

Sébastian Konieczny and Ramón Pino Perez. Mergin information under constraints: A logical framework. *Journal of Logic and Computation*, 12(5): 773–808, 2002.

Christian List. A possibility theorem on aggregation over multiple interconnected propositions. *Mathematical Social Sciences*, 45(1):1–13, 2003.

Christian List. Group communication and the transformation of judgments: An impossibility result. *The Journal of Political Philosophy*, 19(1):1–27, 2011.

Christian List. The theory of judgment aggregation: An introductory review. *Synthese*, 187(1):179–207, 2012.

Christian List and Philip Pettit. Aggregating sets of judgments: An impossibilty result. *Economics and Philosophy*, 18(1):89–110, 2002.

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

David Miller. Deliberative democracy and social choice. *Political Studies*, 40 (1):54–67, 1992.

Klaus Nehring and Clemens Puppe. Strategy-proof social choice on single-peaked domains: possibility, impossibility and the space between. *Department of Economics, University of California at Davis*, 2002.

Nicolas Schwind, Katsumi Inoue, Gauvain Bourgne, Sébastian Konieczny, and Pierre Marquis. Belief revision games. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence*, 2015.

Nicolas Schwind, Katsumi Inoue, Gauvain Bourgne, Sébastian Konieczny, and Pierre Marquis. Is promoting beliefs useful to make them accepted in networks of agents? In *Proceedings of the 25th International Joint Conference on Artificial Intelligence*, 2016.

Sunil Simon and Krzysztof R Apt. Social network games. *Journal of Logic and Computation*, 2015.

Keith Werner and Carl Wagner. *Rational Consensus in Science and Society.* Springer, 1981.