UNIVERSITY OF AMSTERDAM

MASTERS THESIS

Modelling Meta-Agreement through an Agent-Based Model

Examiner:
Dr. Fernando P. Santos

Author:
Amir SAHRANI

Supervisor:
Prof. Dr. Ulle Endriss

Assessor:
Dr. Davide Grossi

A thesis submitted in partial fulfilment of the requirements
for the degree of Master of Science in Computational Science

in the

Computational Science Lab
Informatics Institute

February 2025

Declaration of Authorship

I, Amir SAHRANI, declare that this thesis, entitled 'Modelling Meta-Agreement through an Agent-Based Model' and the work presented in it are my own. I confirm that:

- ☐ This work was done wholly or mainly while in candidature for a research degree at the University of Amsterdam.
- ☐ Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- ☐ Where I have consulted the published work of others, this is always clearly attributed.
- ☐ Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- ☐ I have acknowledged all main sources of help.
- ☐ Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Your signature

Date: February 2025

"The majority, standing in for the people, wills everything and therefore wills nothing"

Joshua Cohen

Abstract

Include your abstract here Abstracts must include sufficient information for reviewers to judge the nature and significance of the topic, the adequacy of the investigative strategy, the nature of the results, and the conclusions. The abstract should summarize the substantive results of the work and not merely list topics to be discussed.

Length 200-400 words.

Acknowledgements

Thank the people that have helped, supervisors family etc.

List of Algorithms

ABBREVIATIONS

CSL   **C**omputational **S**ceince **L**ab

UvA   **U**niversitiet **v**an **A**msterdam

N The set of all voters

X The set of all alternatives

$\succ$ A preference relation ship

$\mathcal{D}$ A domain of possible profiles

$\lhd$ A geometric order over candidates

CHAPTER 1

INTRODUCTION

Though Black's result [1] is a famous positive result, it is far from the only positive result relating to domain restrictions. We first look into various domain restrictions and their properties. To fully understand why single-peakedness is specifically desirable we outline the political and philosophical reasons first, after which we elaborate on the mechanism through which deliberation should result in single-peakedness according to List [2]. Finally, for completeness' sake, we mention critiques of this theory.

We first proceed by giving a short introduction of social choice. We outline the basic model, and restate well known results relevant to the following chapters.

2.1 The basic model

To model elections, or more generally voting games, we represent voters by the set N consisting of n voters. The possible outcomes of an election, we represent with the set A consisting of $|A|$ possible outcomes, from now on we will refer to the outcomes of an election as alternatives. Each voter can represent their preference on alternatives through a preference relation $\succeq_i$, for example if voter 2 prefers outcome a to outcome b, we write $a \succeq_i b$. If, however, a voter wants to make it clear a is strictly better than b, we instead write $a \succ_i b$, when a voter specifies their preferences on the entire set of alternatives we call this a (weak) linear order. We call the set of possible linear orders over the alternatives $\mathcal{L}(A)$, the set of weak linear orders is denoted by $\hat{\mathcal{L}}(A)$. Thus, for an election, all voters report a (weak) linear order, the set of each voters preference is called a profile, denoted by $\mathbf{R}$. Finally, a rule F decides the outcome of the election based on the profile. We discuss the specifics of these rules in section section 2.2.

The last general tool we will need is the *majority relationship*. Given some profile $\mathbf{R}$ can construct a majority relationship as follows, for each pair of alternatives x, y, we ask how many people prefer x to y, if the number of people who prefer x to y is greater than the other way around we write $x \succ_{\text{maj}} y$, if we have an even number of voters, and thus $|N_{x \succ y}| = |N_{y \succ x}|$ this becomes a weak preference, and we simply write $x \succeq_{\text{maj}} y$ (defaulting to lexicographical order). We proceed with an example.

EXAMPLE 1: *Majority relation*

1	2	3
a	b	a
b	c	c
c	a	b

Given the profile on the left, we first start by comparing a to b, both voters 1 and 3 prefer a to b, thus the majority has prefers a to b. Comparing b to c,the majority prefers b to c. Finally, comparing a to c, a is preferred again. Thus, the majority relation is $a \succ_{\text{maj}} b \succ_{\text{maj}} c$

A majority relation can be a-cyclic, and transitive, though neither are guaranteed. An a-cyclic majority profile is simply a majority relation without any cycles, meaning there does not exists a series of alternatives $a_1, \ldots a_n$ such that $a_1 \succ a_2 \succ \ldots \succ a_n \succcurlyeq a_1$. Transitivity is very similar, stating that the preferences between alternatives are transitive in that for any triplet of alternatives x, y, z if $x \succcurlyeq y$ and $y \succcurlyeq z$ then $x \succcurlyeq z$. These notions are similar, but transitivity is a stronger requirement, as it includes indifference.

2.2 Social Choice Functions

In order to decide the outcome of an election, we pick a social choice function F, this function should map all possible profiles to an outcome, thus $F : \mathbf{R} \mapsto A$. A famous example of a SCF is the plurality rule, which simply elects the alternative voted into first place most often, though simple, it can also lead to a tie. Since the outcome of our SCF is only allowed to be a single alternative, the plurality rule needs to be equipped with a tie breaking mechanism in order to be a valid SCF, we require the tie-breaking to be deterministic.

2.2.1 Axioms

Though any voting rule that outputs one alternative for each profile is valid, for real elections organizers likely will want to ensure the rule has certain nice properties, such as not favoring some alternatives of others. In social choice these properties are called axioms, and the procedure of designing a rule based on axioms is called the axiomatic approach. The name of the property just described is the axiom of neutrality, stating that the voting rule should be neutral with respect to the alternatives. In this work three main axioms are of importance.

Axiom of Surjectivity. A SCF F is surjective, if for every alternative, there exists a profile R such that $F(R)$ elects it.

Axiom of Non-Dictatorship. A SCF F is non-dicatorial, if there does not exist a voter i such that $F(R) = \text{top}(i, R)$ for all profiles R, where $\text{top}(i, R)$ extracts voter i's most preferred alternative from profile R.

Axiom of Strategyproofness. A SCF F is strategy proof if, for any voter $i \in N$, i cannot report an untruthful preference, and thereby cause the outcome of the elective to improve for them.

Another way to interpret strategyproofness is that the SCF should maximize the outcome for all voters, as such if a voter reports something which is not their true preference, the outcome will maximize the wrong preference and thus result in an outcome that is worse for you.

There are many more axioms on could reasonably argue for, however, these are enough to lead to the main impossibility result this works focuses on.

2.3   Negative results

Classic social choice theory has many negative results, one such example is the Condorcet cycle. This is a specific profile that results in a cycle in the majority relation, as shown in the following example.

EXAMPLE 2: *Condorcet cycle*

1	2	3
a	b	c
b	c	a
c	a	b

Voters 1 and 3 prefer a to b, forming a majority, next voters 1 and 2 prefer b to c, forming another majority. However, voters 2 and 3 prefer c to a forming a majority, and thus creating a cycle.

It is not hard to convinces oneself that under weak preferences the Condorcet cycle can occur anytime there are 3 or more alternatives and voters. While under strict preferences this can occur anytime the number of alternatives is odd and greater than 3, with the number of voters being a multiple of the number of alternatives. As we will show later, this profile can be the cause of some impossibility results.

One of the major negative results in social choice is that of the Gibbard Satherswaite theorem [3, 4].

Theorem 2.1 (Gibbard-Satherswaite). *There exists no resolute Social Choice Function for elections with $|A| \geq 3$ that is surjective, strategyproof, and non-dictatorial.*

Proof. ...
□

2.4 Domain Restrictions

Many negative results are a consequence of a few ill-behaved profiles, if one can argue such profiles do not occur in the real election, there is some hope of constructing SCF's satisfying our axioms. To speak more formally about profiles "not occuring", we introduce Domain restrictions, for this we use the definition by Elkind et al. [5].

DEFINITION 1: *Domain*

Given a set of voters N, alternatives A, and conditions C, the domain $\mathcal{D}$ of an election is the set of all profiles R such that all conditions C are satisfied.

This definition is different from usual definitions in social choice in so far as it talks about allowed profiles instead of allowed votes.

As stated earlier, the Condorcet profile is one such ill-behaved profile, as each alternative, holds a majority preference over another alternative. Naturally one might consider if this profile might even come up in practice, since though conceivable it seems generally unlikely that there exists a perfect split in opinions. Quite naturally one of the first "solutions" one might consider is when the number of voters is not a multiple of the number of alternatives, though this is hardly a useful solution since it only prevents Condorcet cycles, it is the first example of a domain restriction, we define it as follows

---

DEFINITION 2: $\mathcal{D}_{\text{No-tie}}$

Let X be the set of alternatives and N be the set of voters, of size n such that $n \neq k \cdot |X|$ for any $k \in \mathbb{N}$. We call this domain $\mathcal{D}_{\text{No-tie}}$.

This allows us to state our first proposition.

Proposition 2.2. *The plurality rule never returns a $|X|$-way tie between alternatives when applied to $\mathcal{D}_{\text{No-tie}}$*

Proof. Assume, for the sake of contradiction, the plurality in fact does return a tie this must mean that all alternatives were ranked first an equal number of times, call this k, necessarily then, we have need exactly $k \cdot |X|$ voters, but this leads to a contradiction, as this would no longer be inside $\mathcal{D}_{\text{No-tie}}$. $\qquad \square$

This is a simple result, but it serves as an example on how we can use the properties of the domain to prove things about the election. Gaertner [6] establishes 2 ways in which a domain can be restricted. Firstly we can restrict the domain to a number of voters or alternatives, which is what we did in $\mathcal{D}_{\text{No-tie}}$. Secondly, the domain can be restricted to have a certain structure, such as being single-peaked.

2.4.1   Single-Peaked profiles

In a election the alternatives might represent a axis, such that a voters is prefers an alternative more if they are closer to them on the axis. For example, if the alternatives represent free-trade vs regulation, we can imagine that a voter that is of the opinion that free trade is of ultimate importance will prefer alternatives more the more the are on the side of free trade. More generally, we call a profile single peaked if there exists an axis on which we can place the alternative such that all voters' preferences have a single peak on this axis. Definition 3 makes this notion formal.

DEFINITION 3: *Single-Peaked Profiles*

A profile P is single-peaked, if given some ordering $\lhd$ over the alternatives, it holds that for all voters i, and all $a, b, c \in X$, if $a \lhd b \lhd c$, then either $a \succ_i b$ or $c \succ_i b$, but never both.

3.0.1 Condorcet Domain

If our goal is to prevent Condorcet cycles, or in general have transitive majority relations, the best we could hope to do is to apply our domain restriction such that our domain contains all profiles P such that P has a (weak) Condorcet winner. We call this domain $\mathcal{D}_{\text{Condorcet}}$. Under this domain, let $f_{Condorcet}$ be the Condorcet Rule, which picks a Condorcet winner. Then $f_{Condorcet}$ is strategyproof over $\mathcal{D}_{\text{Condorcet}}$ [5].

Proof. (Elkind et al. [5]). Assume, for the sake of a contradiction, we have profiles $P = (\succ_1 \ldots \succ_i \ldots \succ_n)$ and $P' = (\succ_1 \ldots \succ_{i'} \ldots \succ_n)$ such that:

$$f_{Condorcet}(P) = a, \quad f_{Condorcet}(P') = b, \quad \text{and } a \neq b$$

Then under P a strict majority $N' \subseteq N$ have $a \succ b$, but $i \notin N'$, thus in P', N' is still a majority preferring a to b, but this is in contradiction to b winning in P'. □

Though this result is positive, $\mathcal{D}_{\text{Condorcet}}$ is not hereditary, this is easy to see through an example:

EXAMPLE 3: $\mathcal{D}_{\text{Condorcet}}$ *is not hereditary*

v_1	v_2	v_3	v_4
a	b	c	a
b	c	a	c
c	a	b	b

We can see that in this example, a is the weak Condorcet winner, as it beats b and is tied with c, however if we remove voter 4, we return to the original Condorcet cycle.

A domain not being hereditary means that the nice properties of the domain can be unstable, as the number of voters and alternatives might not be known or could be manipulated. Instead, we might want to look at hereditary domains.

3.0.2 Hereditary Domains

DEFINITION 4: *Hereditary* (Elkind et al. [5])

A domain restriction onto $\mathcal{D}$ is *hereditary* if, for every profile $P \in \mathcal{D}$, and every profile P', that can be obtained by deleting voters and alternatives from P, P' is also in $\mathcal{D}$

The first hereditary domain we present, will also be the main focus of this thesis. This is the domain of all single-peaked profiles.

Note that in a voter is allowed to prefer b best, and then choose a or c in any order. It is clear to see that if any voter or alternative is deleted, this property is satisfied.

Proposition 3.1. (Elkind et al. [5]). $\mathcal{D}_{\text{SP}}$ *is hereditary.*

Proof. (Voter Deletion). If we remove a voter, this does not affect the other voters, thus the property is satisfied. ✓

(Alternative Deletion). Consider a voter i and their single-peaked vote, if we remove some alternative x, this voter all alternatives which voter i preferred to x stay in the same position, while all other alternatives move up one rank, thus preserving the order. ✓ □

A similar notion to single-peaked profiles is that of single caved profiles, which is equivalent, but instead a voters peak representing their most preferred option, they have a valley, which represents the worst option. Single caved profile are hereditary as well, but for a voting rule on them to be strategy proof, only two possible alternatives can be chosen, the left and right most alternatives according to ◁.

Instead of ordering the alternatives, we can imagine instead ordering the voters, such that we have a leftmost and rightmost voters, and all other voters can be placed between them based on their difference. In this case, a profile is single-crossing if, for any alternative a, its preference relation to another any alternative b flips at most once when traversing the voters in order $\lhd$.

DEFINITION 5: *Single-Crossing Profiles* (Elkind et al. [5])

A profile P is single-crossing w.r.t. some ordering $\lhd$, if for any $a, b \in X$, $\{i \in N : a \succ b\}$ and $\{i \in N : b \succ a\}$ are both intervals over $[n]$. A profile P is single crossing if the votes can be permuted such that it is single crossing w.r.t. a given ordering.

Similar to single-peaked profiles, the domain of single-crossing profiles, $\mathcal{D}_{\mathrm{SC}}$ is also hereditary

Proposition 3.2. $\mathcal{D}_{\mathrm{SC}}$ *is hereditary*

Proof. (Voter Deletion). Deleting a voter preserves the ordering between voters, as such this cannot introduce a new crossing between alternatives. ✓

(Alternative Deletion). If we remove an alternative the voters' rankings of the other alternatives does not change, thus preserving single-crossing. ✓ □

3.1 The History of Deliberation and Meta-Agreement

We have provided an overview of different domain restriction and their properties, mainly showing how they avoid Condorcet cycles. Some argue however, that Condorcet cycles are empirically rare. The next section is dedicated to explaining why this is so through examining the historical ideas around deliberation and deliberative democracy, as well as that of Meta-Agreement.

3.1.1 Deliberation

Though deliberation is intuitively familiar, namely the process of multiple people talking through a problem with the goal of coming to an agreement, compromise or solution, providing a definition that is both clear and consistent with the literature in Political Science, Philosophy and Social choice is difficult. This intuition it leaves some of the reasons for and goals of deliberation, as state in the literature, unmentioned.

Freeman [7] gives an overview of deliberative democracy, in which he shares the intuitive idea that a deliberative democracy contains open discussion, open legislative deliberation

and a pursuit of the common good. He also notes that there is no common agreement on the central features of a deliberative democracy, one account is that of deliberative democracy simply involving discussion among the public before voting. Another similar account is that this voting must not only be preceded by deliberation, but also general communication, all of which intended to change people's preferences. He further proceeds to give a more detailed conception of deliberative democracy, according to which a deliberative democracy is one in which political agents or their representatives

1. Aim to collect, deliberate and vote

2. Represent their sincere and informed judgements

3. Vote and deliberate on measures beneficial to the common good on the citizens

4. Are seen and see each other as political equals

5. Have Constitutional right and social means enables them to participate in public life

6. Are individually free, such that they have their own freely determined conceptions of the good

7. Have diverse and disagreeing conceptions of the good

8. Recognize and accept their duty as democratic citizens, and do not engage in public argument on the basis of their particular moral views incompatible with public reason.

9. Agree reason is public, in so much as it is related to and advances common interests of citizens

10. Agree that their common interest lies primarily in freedom, independence and equal status as citizens.

These features allow us to be more precise when we talk about a deliberative democracy, and in turn be more careful about what deliberation must entail. Cohen [8] further argues that deliberation is needed for democratic legitimacy. By this he means that without deliberation, a democracy is simply the will of the majority, but since majority rule is unstable, it is simply a reflection of the particular institutional constrains at the time. He further goes on to describe the *ideal deliberative procedure* as follows

1. Ideal deliberation is *free*, participants regard themselves as only bound by the results of the deliberation, and the preconditions thereof. Participants act in accordance with the decision made through deliberation, and it being agreed on is sufficient reason to do so.

2. Ideal deliberation is *reasoned*, parties are required to state their reasons for advancing proposals.

3. In ideal deliberation, parties are *equal*, both formally and substantively. There are no rules that single individuals out, and existing distributions of power to no lend a party the opportunity to contribute to deliberation.

4. Ideal deliberation aims to arrive at *consensus*, which can be rationally defended.

3.1.2   Meta-Agreement

A goal of deliberation could be to reach consensus, which is sometimes referred to as substantive agreement, Elster [9] argues that this is not only the goal, but through unanimity this process completely replaces voting, thereby circumventing Arrow's impossibility theorem: "Or rather, there would not be any need for an aggregation mechanism, since a rational discussion would tend to produce unanimous preferences." (p. 112). Though it would be desirable to circumvent Arrow's impossibility theorem, in practice people, even after deliberation might not, indeed often do not, come to full substantive agreement. List [2] instead proposed another lens through which we can analyze deliberation and the type of agreement it induces.

Under *Meta-agreement* individuals do not need to agree on their most preferred outcome, instead they only need to agree on the dimensions of the problem. To contrast this with substantive agreement, under which individuals do not need to conceive of the problem in the same way, all they need is to agree on the same outcome. This means that under substantive agreement, voters can agree outcome $a \succ b$ for different reasons, while under meta-agreement, if voters disagree on $a \succ b$ it must be for the same reason.

According to List [2] there are three hypotheses that need to be satisfied for deliberation to induce meta-agreement:

D1 Deliberation leads people to discover a single *issue*-dimension

D2 Deliberation lets people place all possible alternatives in this *issue*-dimension

D3 After this deliberation, people update their preferences by picking a preferred (peak) outcome, and all other rankings are based on structure of the *issue*-dimension

All these are necessary conditions for meta-agreement, from this is it also clear to see that, given that there is exactly 1 *issue*-dimension, single-peaked profile are, by definition, a direct consequence. This is the main reason meta-agreement is desirable, as it lets us circumvent the Gibbard-Satterthwaite theorem [3, 4] through restricting the domain of preference profiles to the single-peaked domain $\mathcal{D}_{\mathrm{SP}}$

List et al. [10] provide empirical evidence for this theory of deliberation, showing deliberation increases proximity to single-peakedness, which they define as $S = \frac{m}{n}$ where $n = |N|$ and m is the largest subset of voters such that their profile is single-peaked.

Furthermore, they also introduce the notion of salience, which represents to what extent a topic is salient in the voting population. In order to test whether deliberation increases single-peakedness *through* meta-agreement, they test the following four hypotheses: (H1) deliberation increases proximity to single-peakedness. (H2')[1] high salience issues show less increase in PtS then low salience issues. (H3) Effective deliberation, in the sense that more is learned during deliberation, results in bigger increases. (H4) All things equal, the increase is largest for issues with natural *issue*-dimensions.

Meta-agreement is not without its critiques, however. Ottonelli and Porello [11] show meta-agreement to be a stronger requirement than it may seem at a first glance. Firstly for (D1) to hold, the *issue*-dimension must hold some semantic meaning, as otherwise it is unclear how people can exchange conceptualization of the problem otherwise. Furthermore, the issues must consist of 2 semantic issues, otherwise with only 1 dimension voters simply reach substantive agreement. A further restriction on these two dimensions is that they need to be opposite, with opposite justifications. If this is not the case, a voter can agree with both justifications, and thereby introduce a new dimension "balance", which then violates the conditions under which single-peaked profiles guarantee the existence of strategyproof voting rules. D2 requires that all voters share the exact same semantic understand of the dimension, and the outcome associated with each alternative. Finally D3 requires D1 and D2 to have happened before in order. Clearly D3 is the weakest of the three.

Thus, meta-agreement is still quite restrictive, needing multiple forms of unanimity, and only applying to problems with certain properties. Nonetheless, in this work we investigate its explanatory power on prevention of Condorcet cycles.

3.2 Related Work

Rad and Roy [12] model deliberation and its effect on single-peakedness, though they argue single plateauedness is a more accurate term. To this end, they model each voter to have preferences order, and deliberation being the process of all voters announcing their preferences, after which all other voters update their current preference towards that of the announced ranking, in doing so they might have a bias towards their own preference, as such they try to minimize the distance between their current preference and the announced one. This process repeats until all voters have announced their opinion once, for one or multiple rounds. The preference a voter adopts when updating must lie between their current profile and the announced profile, which profiles are considered to be "between" is defined by the distance metric used. They considered

---

[1]This is a test for a corollary. H2 states that the rate of increase of proximity to single-peakedness decreases. Since high salience means some sort of deliberation has happened before, we expect this to have the same affect.

three metrics, the Kemeny-Snell (KS) [13], Duddy-Piggins (DP) [14], and Cook-Seiford (CS) [15]. Both KS and DP depend on the judgement set resulting from the voters preferences, the KS distance is then defined as the number of binary swaps a judgement set needs to undergo before it becomes the target judgement set, an example for such a swap would be going from $(a \succ b)$ to $\neg(a \succ b)$. The DP distance is defined on the graph of judgement sets, where 2 sets share an edge if there is no judgement set between them. Since KS and DP share their notion of betweenness, we introduce betweenness as follows.

DEFINITION 6: *J-Betweenness*

A judgement set J_i is between profiles J_j and J_k if for every proposition about $x, y \in A$, J_i either agrees with J_j or J_k.

Figure 3.1 shows a graph used for the DP distance in the case of 3 alternatives, for simplicity the associated profiles are used to label the judgement sets.

The CS distance is simpler and is simply defined as the number of positions two voters disagree on, and a profile is between two others if for each position it agrees with one of the two profiles.

Each distance has different trade-offs, CS is the simplest, but might exaggerate the distance when there are many alternatives, for example if 2 voters agree on the relative ranking of all but 1 alternative, which one voter happens to rank first, thereby shifting all other profiles right. The KS distance, using judgement sets instead of raw profiles captures this more effectively, while still being relatively easy to compute, but in cases of more disagreement, it is likely to over count the distance, since the binary changes to not capture logical necessities. For example, swapping $(a \succ b)$ to $\neg(a \succ b)$ must result in $(a \succ b)$ becoming true (in the case of strict preferences), thus one might reasonably conclude this should only count as 1 step. DP improves upon this, but in doing so becomes much harder to compute, mainly through the cost of constructing the full graph of judgement sets, which grows in $\mathcal{O}(n!)$ in the number of vertices, where n is the number of alternatives.

Apart from there distances, they also define a voter as a tuple of a (weak) preference and a bias (towards their current position) $v = \langle r, b \rangle$, with $b \in \mathbb{R}_{[0,1]}$. Finally, a deliberation step $D_s : V \times r \mapsto V$, with V being a set of voters and s being one of the spaces (KS, DP, CS). A round of deliberation consists of n deliberation steps, where each voter has announced their opinion once. We formulate this procedure in the following program:
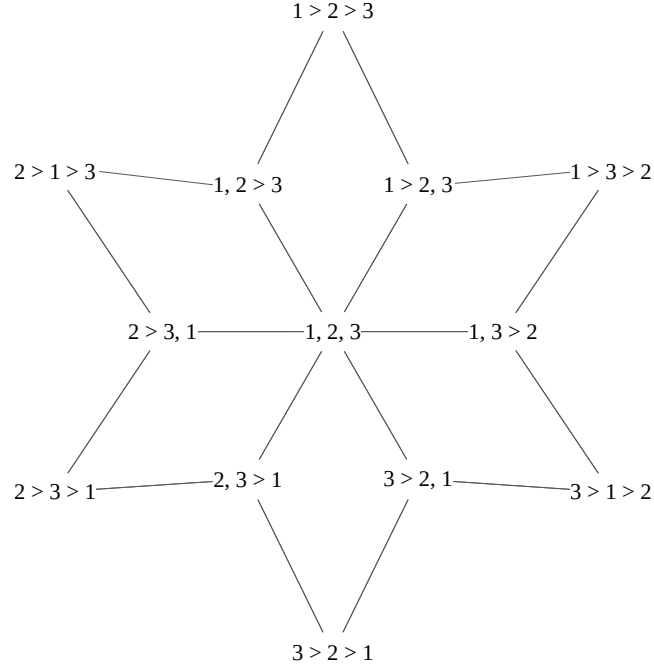
FIGURE 3.1: The graph of judgement sets for all preferences over three alternatives.

---

input : Set of Voters V, metric space s
output: Updated set of Voters V

$V_{\mathrm{u}} \leftarrow V$ `// Set of unannounced voters (references to V)`
while $|V_u| > 0$ **do**
    Select a random $v \in V_{\mathrm{u}}$
    $V_{\mathrm{u}} \leftarrow V_{\mathrm{u}} \setminus \{v\}$
    $V \leftarrow D_s(V, v.r)$ `// Update voters based on v's preference`

The deliberation step D_s then updates all voters such that their new preference minimizes the following formula.

$$\sqrt{bd_s(r_i, r')^2 + (1-b)d_s(r_j, r')^2} \tag{3.1}$$

Where r_i, r_j are the voters and the announced preference, respectively, and r' is the voters new preference.

We present a replication and extension of their work chapter 6. Furthermore, we present novel results based on this model in chapter 4.

THEORETICAL RESULTS

The following are propositions are not fully worked out, which is also why they are written informally, but general ideas I got from toying around with the model. I also have not spent substantial time thinking how to prove them. The point of the first two is to show that deliberation itself is not strategyproof, at least not in this form. I am not sure if whatever model of deliberation we end up creating will be (although I am highly doubtful that it will). The point is also not to say that it needs to be strategyproof, we can simply argue that deliberation not being strategyproof is preferable as voters at least have a change of recognizing some point lying etc.

In the model of deliberation of Rad and Roy [12], outlined in section 3.2, aims to model deliberation, through this the aim is to show deliberation results in nicely structured profiles which allow for strategy proof voting rules. One important caveat, also given by the authors, is that all participants should honestly and truthfully participate in deliberation. We attempt to make this more formal with the following proposition.

Proposition 4.1. *Deliberation on any metric space can be manipulated if all voters can still change their mind (open-minded voters), if one voter decided not to change their mind.*

Proof. (very rough sketch). If all voters are able to change their mind, a malicious voter, call them $\hat{v}$, can simply refuse to change their mind. Since all other are still updating their preferences, if they currently hold a preference that still has a "path" towards $\hat{v}$'s preference they will end start to agree. $\square$

The two main issues I have when I think about actually writing a proof is that the hole notion of open-minded voters is a bit vague (e.g. 0.73 is just and empiric observation), thus it would need some kind of lemma showing when voters stop being open-minded. Furthermore, I have trouble thinking about making formal statements about a deliberation step in the first place

Unfortunately, sticking to your original preference is not the only way in which a voter can manipulate the outcome of a deliberation. Through misrepresentation a voter could argue for a preference different from their in the hope to better the outcome. As an example imagine after 1 round of deliberation voters find themselves separated into two clusters, with mutually exclusive opinions, as in given the metric space there are no opinions between the two clusters. A strategic voter could then pretend to hold an opinion that is compatible with the other clusters and slowly move them towards their

Proposition 4.2. *Deliberation on any metric space can be manipulated if all voters can still change their mind, through misrepresentation of one current preferences.*

I have similar issues with this one, but I think this could go two ways, either we don't let the strategic voter change their main (bias = 1), which then would build on the previous proposition. Or instead the voter can change their mind, but they will present a false preference in their announcement. I think the second approach might result in a possible complexity theory result.

Then from a game theoretic perspective, we can model deliberation as a game, through which voters try to manipulate the outcome of the election.

CHAPTER 5

METHODS

We determine whether our profiles are single-peaked using the algorithm by ..., implemented in `Ocaml` as follows

```ocaml
let agent = 1 in
let text = "Testing" in
(* This is a comment *)
print_int agent
```

EXPERIMENTAL RESULTS

We first present a full replication and extension of the work by Rad and Roy [12]. Then we present the simulations based on our model of meta-deliberation, as well as the results of the sensitivity analysis on both models.

6.1 Replication

We are able to fully replicate the results found by Rad and Roy [12], in figure 6.1 we see while the bias is less than 0.73 all metric results in a-cyclic preferences. We also replicate the behavior of the KS-metric with biases in the range of 0.73-0.85, showing that it causes even some a-cyclic profiles to become cyclic. Figure 6.2 shows a similar graph, but instead of the number of unique preference orders present after deliberation. Rad and Roy [12] refer to these as clusters. As we can clearly see, when bias is below 0.5, voters always agree on 1 preference order, between 0.5 and 0.73, they agree on 2 preferences for all distances. After this DP and CS both show that there is complete diversity of opinion, with all preferences being present, while KS first shows that between 0.73 and 0.85 the are 3 preference orders after deliberation, which also happen to be cyclic (we can draw this conclusion based on the fact that for this range KS introduces more cycles). Finally, after 0.85 KS behaves like DP and CS, as no voter changes their mind anymore.
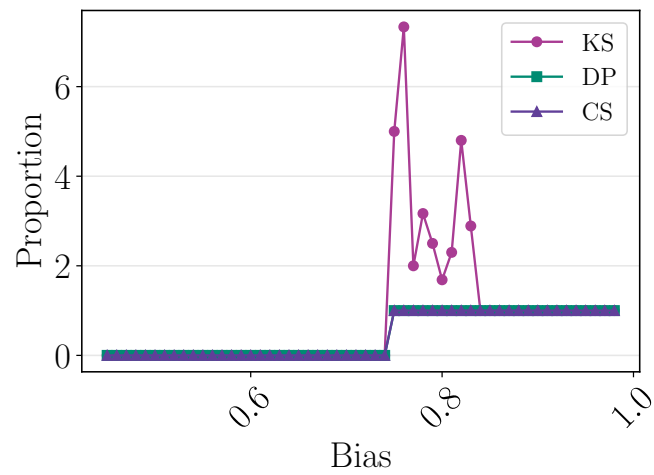
FIGURE 6.1: The proportion of cyclic profiles remaining, 0 indicating that no cyclic profiles were present after deliberation.
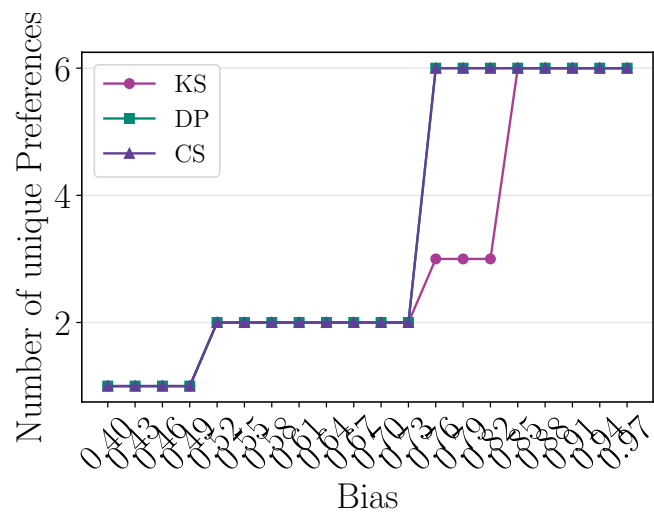

FIGURE 6.2: The proportion of intransitive profiles remaining, 0 indicating that no cyclic profiles were present after deliberation.

CHAPTER 7

DISCUSSION

CHAPTER 8

Conclusion and future work

Ethics and Data Management

A new requirement for the thesis is that there must be a short section in which you reflect on the ethical aspects of your project. This requirement is related to one of the final objectives that a graduated student of the Master of Computational Science must meet: "The graduate of the program has insight into the social significance of Computational Science and the responsibilities of experts in this field within science and in society". You don't need to devote an entire chapter to this; a short section or paragraph is sufficient.

I acknowledge that the thesis adheres to the ethical code (https://student.uva.nl/en/topics/ethics-in-research) and research data management policies (https://rdm.uva.nl/en) of UvA and IvI.

The following table lists the data used in this thesis (including source codes). I confirm that the list is complete and the listed data are sufficient to reproduce the results of the thesis. If a prohibitive non-disclosure agreement is in effect at the time of submission "NDA" is written under "Availability" and "License" for the concerned data items.

Short description (max. 10 words)	Availability (e.g., URL, DOI)	License (e.g., MIT, GPL, CC)
Example dataset 1	\<github url\>or Figshare	GPL
Example source code	DOI (from Zenodo)	MIT
Example sensitive data	NDA	NDA

Bibliography

[1] Duncan Black. On the Rationale of Group Decision-making. 56(1):23–34. ISSN 0022-3808. doi: 10.1086/256633. URL https://www.journals.uchicago.edu/doi/abs/10.1086/256633.

[2] Christian List. Two Concepts of Agreement. 11(1):72–79. ISSN 1538-9731. URL https://muse.jhu.edu/pub/2/article/12234.

[3] Allan Gibbard. Manipulation of Voting Schemes: A General Result. 41(4):587–601. ISSN 0012-9682. doi: 10.2307/1914083. URL https://www.jstor.org/stable/1914083.

[4] Mark Allen Satterthwaite. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. 10 (2):187–217. ISSN 0022-0531. doi: 10.1016/0022-0531(75)90050-2. URL https://www.sciencedirect.com/science/article/pii/0022053175900502.

[5] Edith Elkind, Martin Lackner, and Dominik Peters. Preference Restrictions in Computational Social Choice: A Survey. URL http://arxiv.org/abs/2205.09092.

[6] Wulf Gaertner. Chapter 3 Domain restrictions. In *Handbook of Social Choice and Welfare*, volume 1 of *Handbook of Social Choice and Welfare*, pages 131–170. Elsevier. doi: 10.1016/S1574-0110(02)80007-8. URL https://www.sciencedirect.com/science/article/pii/S1574011002800078.

[7] Samuel Freeman. Deliberative Democracy: A Sympathetic Comment. 29(4): 371–418. ISSN 1088-4963. doi: 10.1111/j.1088-4963.2000.00371.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1088-4963.2000.00371.x.

[8] Joshua Cohen. DELIBERATION AND DEMOCRATIC LEGITIMACY. In *Debates in Contemporary Political Philosophy*. Routledge. ISBN 978-0-203-98682-0.

[9] Jon Elster. THE MARKET AND THE FORUM: Three varieties of political theory. In *Debates in Contemporary Political Philosophy*. Routledge. ISBN 978-0-203-98682-0.

[10] Christian List, Robert C. Luskin, James S. Fishkin, and Iain McLean. Deliberation, Single-Peakedness, and the Possibility of Meaningful Democracy: Evidence from Deliberative Polls. 75(1):80–95. ISSN 0022-3816, 1468-2508. doi: 10.1017/S0022381612000886. URL https://www.journals.uchicago.edu/doi/10.1017/S0022381612000886.

[11] Valeria Ottonelli and Daniele Porello. On the elusive notion of meta-agreement. 12(1):68–92. ISSN 1470-594X. doi: 10.1177/1470594X11433742. URL https://doi.org/10.1177/1470594X11433742.

[12] Soroush Rafiee Rad and Olivier Roy. Deliberation, Single-Peakedness, and Coherent Aggregation. 115(2):629–648. ISSN 0003-0554, 1537-5943. doi: 10.1017/S0003055420001045. URL https://www.cambridge.org/core/journals/american-political-science-review/article/deliberation-singlepeakedness-and-coherent-aggregation/652CCF8DF5B03EF352B4B04AD66D90C0.

[13] John G Kemeny and given-i=LJ family=Snell, given=LJ. Preference ranking: An axiomatic approach. pages 9–23.

[14] Conal Duddy and Ashley Piggins. A measure of distance between judgment sets. 39(4):855–867. ISSN 0176-1714. URL https://www.jstor.org/stable/41681743.

[15] Wade D. Cook and Lawrence M. Seiford. Priority Ranking and Consensus Formation. 24(16):1721–1732. ISSN 0025-1909. doi: 10.1287/mnsc.24.16.1721. URL https://pubsonline.informs.org/doi/abs/10.1287/mnsc.24.16.1721.