Anchoring in deliberations of structured groups

Inaugural-Dissertation

zur Erlangung des Doktorgrades der Philosophie
der Ludwig-Maximilians-Universität München

vorgelegt von

José Antonio Leyva Galano

aus
Holguín (Kuba)

2024

Referent: Prof. Dr. Stephan Hartmann
Korreferent: apl. Prof. Dr. Martin Rechenauer
Tag der mündlichen Prüfung: 23. Februar 2024

Abstract

Group deliberation is the central topic of this dissertation. Specifically, here we study the relevance of the order of the speakers to deliberations that happen in groups.

People deliberate in groups all the time. They deliberate at home, in governmental institutions or at their working place for deciding on simple or less simple matters. Naturally, the outcome of a group deliberation might be influenced by some expected factors: the different options that will be considered, the rationality of the speakers, and the accepted ways of argumentation. However, it would be unfortunate if the outcome of a deliberation was influenced by the order in which the speakers presented their arguments. This idea motivates the leading question of this dissertation: does a deliberative situation favour the first speaker of the group?

Our leading question will take slightly different forms through the dissertation, and to be able to answer them we use models as a medium of representation. That is, we take a deliberative situation, we describe it as its model-structure, and we answer our original question on that structure. Likewise, in the case of a similar question, this time regarding a family of deliberative situations instead of a single one, we turn the question and the scenario into a question and a scenario about a family of models.

The central result of this dissertation contrasts the opinion-strength of the first speaker with the opinion-strength of any other individual that takes part in a deliberation. Broadly said, it claims that if we consider a family of possible deliberative situations, the "region" in which the first speaker has advantage is visibly larger than the one in which this does not happen. Moreover, this effect increases with the number of acceptable opinions in a debate.

The previous result should be uncomfortable to any deliberative account that intends to fulfill the next two constraints at the same time: the first constraint requires that the account follows the generic structure of deliberations that we present in this work. The second one requires that from a moral, utilitarian or epistemic perspective, deliberations take place in a fair environment. Consequently, in order to be on the safe side, deliberative accounts under this category should always provide arguments that show them to be immune to this order dependence concern.

Zusammenfassung

Gruppenberatung ist das zentrale Thema dieser Dissertation. Insbesondere untersuchen wir hier die Relevanz der Reihenfolge der Redner für Beratungen, die in Gruppen stattfinden.

Menschen beraten sich ständig in Gruppen. Sie beraten sich zu Hause, in staatlichen Einrichtungen oder an ihrem Arbeitsplatz, um über einfache oder schwierigere Angelegenheiten zu entscheiden. Natürlich kann das Ergebnis einer Gruppenberatung durch einige erwartete Faktoren beeinflusst werden: die verschiedenen Optionen, die in Betracht gezogen werden, die Rationalität der Sprecher und die akzeptierten Argumentationsweisen. Es wäre jedoch bedauerlich, wenn das Ergebnis einer Beratung durch die Reihenfolge, in der die Redner ihre Argumente vortragen, beeinflusst würde. Dieser Gedanke motiviert die Leitfrage dieser Dissertation: Wird in einer Beratungssituation der erste Sprecher der Gruppe bevorzugt?

Unsere Leitfrage wird im Laufe der Dissertation leicht unterschiedliche Formen annehmen, und um sie beantworten zu können, verwenden wir Modelle als Darstellungsmittel. Das heißt, wir nehmen eine Beratungssituation, beschreiben sie als ihre Modellstruktur und beantworten unsere ursprüngliche Frage anhand dieser Struktur. Im Falle einer ähnlichen Frage, die sich auf eine Familie von Beratungssituationen statt auf eine einzelne bezieht, ändern wir die Frage und das Szenario in eine Frage und ein Szenario über eine Familie von Modellen.

Das zentrale Ergebnis dieser Arbeit vergleicht die Meinungsstärke des ersten Sprechers mit der Meinungsstärke jeder anderen Person, die an einer Beratung teilnimmt. Grob gesagt behauptet es, dass, wenn wir eine Familie möglicher Beratungssituationen betrachten, die “Region”, in der der erste Sprecher einen Vorteil hat, deutlich größer ist als diejenige, in der dies nicht der Fall ist. Außerdem nimmt dieser Effekt mit der Anzahl der zulässigen Meinungen in einer Debatte zu.

Das obige Ergebnis stellt ein Problem für jeden Beratungsansatz dar, der gleichzeitig die zwei folgenden Bedingungen erfüllen will: Dass er erstens der generischen Struktur von Beratungen folgt, die wir in dieser Arbeit vorstellen und dass zweitens die Überlegungen aus einer moralischen, utilitaristischen oder erkenntnistheoretischen Perspektive in einem fairen Umfeld stattfinden. Um auf der sicheren Seite zu sein, sollten Beratungsansätze dieser Kategorie daher immer Argumente liefern, die zeigen, dass sie gegen diese ordnungsabhängigen Bedenken immun sind.

Acknowledgements

Writing acknowledgements is very pleasant. Today, I am happy to thank Mr. Horn and Mr. Gómez, their constant support during my early years in Germany changed my life. I am deeply grateful to Professor Stephan Hartmann, he has been a terrific supervisor. Stephan, it has been a pleasure to work with you. Many thanks to the big Rogosch-Leyva family for their constant encouragement. More importantly, colorful thanks to Bruno, Toni and Linda for everything.

Contents

Introduction

Group deliberation is the central topic of this dissertation. Specifically, here we study the relevance of the order of the speakers to deliberations that happen in groups.

People deliberate in groups all the time. They deliberate at home, in governmental institutions or at their working place for deciding on simple or less simple matters. For instance, some colleagues might engage in a quick exchange at work for picking a place where to eat lunch today. Clearly, each of them might have a preference for a particular place. After all, the options of cafeterias, canteens or coffee shops could be a colorful legion in their neighborhood. To add some complexity, not only the quality of the meal plays a role in this situation, but there might be inner fights running inside each person too regarding diets, lunch budgets, etc. At this point, one could foresee that with little effort this situation can be further entangled as much as one might require. Luckily for humanity constraints exist, and the colleagues can not exchange and justify their opinions forever because the time for lunch-break is bounded, and hungriness around noon speaks for short deliberations too. Moreover, people use to give a high significance to the time-saving-aspect in their lives, so they might not want to expand extensively on “mundane” topics. Consequently, one might expect that after some exchange a plausible scenario emerges, and either the group goes together to the same place, or alternatively they split in smaller groups, and each of them goes to a different place.

Naturally, deliberative situations are often more complicated than our previous example. Everyday, teams of decision makers need to find agreement in diverse scenarios where “the right decision” is either less than obvious or non-existing at all. For instance, the c-level of a company might need to take a strategic decision regarding the size of the organization. Moreover, some members of the board might share the opinion that “now” is the right moment for actively increasing the number of employees. Consequently, they might argue that the finance of the company looks strong and the sales could escalate by opening new representations in different regions of the country. A second group might take a more conservative stance, they prefer to consolidate the current position for a longer period before making an essentially risky move that might damage the good reputation and strong finance of the organization in the case it failed.

In contrast with our first example in which uncertainties were triggered by the existence of preferences corresponding to different options, in the previous case, uncertainties emerged because of the contingent nature of the situation in question: even if all members of the board prefer "the best" for the organization, their different opinions might be caused by their different sensitivities towards risk, diversity of personalities, different career situations or economic circumstances.

Naturally, further examples can be enumerated: teams in consulting agencies need to decide whether a new promising technology is worth the investment. Juries in court must find that someone is guilty or innocent of an illegal act according to the rules of their society. The representatives of a municipality need decide whether pursuing a wind farm is a correct way of employing a portion of the local territory. Groups of experts are required to agree on thresholds for pollutants, so that their governments guarantee to the citizens the existence of a healthy environment.

In these examples of potential deliberative situations, one can already identify some general ingredients of a deliberation: a group of individuals, some options that these individuals will consider during their discussion, the actual deliberative process and its outcome. These are, let's say, the visible ingredients. But, there are other more subtle components that play an important role as well: first, the preferences of each individual regarding the different options that might have originated a debate. Second, the rationalities of the individuals, which involves at least two factors: the particular way in which each individual assesses the arguments given by others **and** the specific forms in which they update their own opinions along a debate. Third, the personalities of the individuals and their particular mental states during a debate.

These basic elements are at the center of the idea of deliberation, and their mention here is essential because they will be explicitly or implicitly very present in the rest of the chapters. However, in this introductory part, we need first to touch **three** basic factors in order to give some shape to the structure of the dissertation, and we do it next.

First, "Why/How is the study of deliberations important in philosophy?". That is, we have claimed that here we study the relevance of the order of the speakers to deliberations that happen in groups. Consequently, next we would like to describe a broad scenario, where the results of this study are important in philosophy: the term deliberation is deeply connected to that of choice, and from experience, we all know that usually the end point of a deliberation is a choice. Naturally, the possibilities involved in the act of choosing might be tangible, like "a glass of cold water" *vs.* "a glass warm water". But, they might be intangible too, like "believing" *vs.* "not believing" that a certain rumor is true. More importantly, the question about the nature of choice is central in philosophy, and it is inevitable that by extension it touches the notion of deliberation too (see *deliberation* and *choice* in Blackburn 2016). Broadly said, **if we pursued** the benefits of "good" group deliberations (either on moral, epistemic or any other ground) as an instrument for delivering (individual) choice, **and** we were not able to generate those deliberations, one might be entitled to challenge our ability to control the process of (individual) choice. Consequently, this should be a clear reason (i.e., the inevitability of bad deliberations might contribute to the tension between *existence vs. non-existence* of a controller "self" in the process of choice) for approaching the study of deliberations with a genuine philosophical interest.

A second kind of interest in deliberations comes from a slightly different area of philosophy (**and we are more interested in this one**). Broadly said, deliberations are important in philosophy because they are involved in many democratic processes. That is, in the context of political philosophy, the subject of deliberation arises as a key element behind the idea of deliberative democracy (see Bächtiger et al. 2018). For this view of democracy, a low probability of reaching an unbiased deliberation would be problematic because it uses an ideal of "good deliberation" as a crucial normative element (often granted with epistemic properties too). Connected to this idea, there is a result that we present in Chapter 2: with high probability, a potential deliberative situation that might occur in a group favours the first speaker. That is, there is a bias component in a large family of deliberations. Of course, these are not good news for the deliberative cause. But, given the increasing number of practical applications (see Bächtiger et al. 2018, p. 19-43 for a broader description) in which a deliberative approach has been employed to enhance a decision making process (this includes environments of deliberative negotiation, deliberative assessment of catastrophic risks, **and more importantly**, instances of deliberative situations within frameworks of *representative democracies*), one might argue that the interest in the study of deliberations is well founded in a broader sense.

Second, given that later we will present some results on models of deliberation, in the second section of this chapter we start preparing for that by answering some illustrative questions. For instance, there we introduce the form in which deliberative situations are represented in this work. Likewise, we describe under which conditions do the results that we present here apply. Naturally, some of the answers are meant to be informal because in their respective chapters there will be more technical details.

Third, because deliberations are often connected to some degree with voting systems (e.g., deliberating before voting *or* deliberating and then voting upon disagreement), in the third section we present three classic results of *Social Choice Theory* (see List 2022), and we also explain why is that a voting component is not playing a role in the work that we present here.

Finally, before closing this introduction, we will summarize the content of each of the chapters of the dissertation.

Deliberative Situations

In this section, our intention is to present simple answers to the following eight questions: **How** do we approach the study of deliberations in this dissertation? **How** do we actually represent a deliberative situation? **How** could our main result be summarized? Under **which** conditions do the results that we present here apply? **Why** is a new notion of a model of deliberation needed? **How** is the term "anchoring" used here? **Is there** empirical evidence of the existence of the *order dependence* reported in this work? **What** is the novelty value of this dissertation?

First. We study deliberations with the assistance of models. In more detail, a recurring question in this dissertation is "does a deliberative situation favour the first speaker of the group?", and we use models for studying this question. That is, we take a deliberative situation, we describe it as its model-structure, and we answer our original question on that structure. Likewise, in the case of a similar question, this time regarding a family of deliberative situations instead of a single one, we turn the question and the scenario into a question and a scenario about a family of models. Naturally, the model-approach that we have described in this paragraph is standard, and advantages and disadvantages from a philosophical perspective are known (see Frigg and Hartmann 2020).

Second. As we mentioned before, we study deliberations with the assistance of models, and **each** model will be described by using **many** tree-like structures called *opinion tree* (see Figure 1.1 for an early diagram). That is, we will need many opinion trees for the description of a single model, one per each deliberative situation that the model describes. Moreover, in this kind of tree-like structures, all vertices are labeled, and each label describes the opinion of each individual of the group at a certain moment of the deliberation. Additionally, the potential transition between two group-opinions is described by a probability value, which is placed between them.

Third. Our main result contrasts the opinion-strength of the first speaker with the opinion-strength of any other individual that takes part in a deliberation. Said in terms of Figure 1, if we consider a complete family of possible deliberative situations, the "region" (of potential deliberative situations) in which the first speaker has advantage is visible larger than the region in which it does not happen. Additionally, as we mentioned in an earlier section, this circumstance should be uncomfortable to any deliberative account that fulfills at the same time the following two demands: first, it follows the structure of deliberations that we present in this work. Second, it requires that from a moral, utilitarian or epistemic perspective a deliberation takes place in a fair environment. Naturally, intuitive reactions to this challenge might be: first, in "our" debates we use different structures of deliberation, and in these structures, the reported problem does not occur. Second, this result represents an essential problem, but from a pragmatic perspective the impact is not too significant.

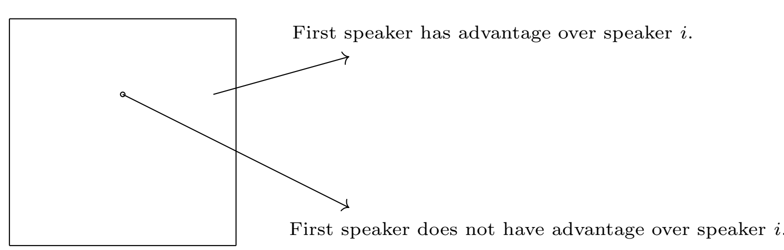

Figure 1: Informal description of the space of deliberative situations.

Fourth. From a formal perspective, the results that we present here are about models that capture a specific (but **very generic**) kind of structure of deliberation. This is an important requirement, and it can not be assumed that a conclusion that holds on those structures can be extended without proof to a different one. Also, most of our (key) results were analytically proved for a setting in which the number of acceptable opinions (in a deliberation) is large. So, under those conditions our results are (analytically) true. We also showed (via simulations) that the same results seem to hold true for some settings with a small number of acceptable opinions. In these cases, we have simulated evidence about the truth of the general statements, but it might be that in some non-simulated instances, the general statements failed to be true. Naturally, from a philosophical perspective, there is also the more general doubt about the adequacy of using models in science. This is a broad question that we should not discuss here, but giving our approach, our position on that regard is completely clear.

Fifth. In this work, the need for a new notion of a model of deliberation has four major reasons. The **first one** is that we conduct a study of deliberations in which the update of the speakers' opinions is asynchronous (i.e., not all agents update their opinions simultaneously). In more detail, we needed a description of deliberation that allowed us to talk about the idea of "the order in which the speakers present their minds", which incidentally triggers an asynchronous pattern for the update of opinions. This requirement already ruled out the standard versions of some prominent instances of models of deliberation (e.g., Lehrer and Wagner 1981 *and* Hegselmann and Krause 2002). A **second reason** is that we are interested in the analysis of bounded deliberations, which are deliberations with a plausible finite number of rounds (i.e., we are not curious here about deliberations that stop or converge on the limit). A **third reason** is that here we confine our attention to debates that allow *uncertain transitions* between two groups of opinions, which is nothing else but accepting the existence of uncertainties in deliberative processes. Unfortunately, when together, the second and the third reasons rule out most common modifications of the Lehrer-Wagner and Hegselmann-Krause models as well. A **fourth reason** is the research method that we have followed in this dissertation. That is, originally we had a strong model-based evidence of the existence of order dependence in a particular model of deliberation (see Hartmann and Rafiee Rad 2020). Accordingly, motivated by this case, we wanted to investigate whether the same remained true for other alternatives of models of deliberation. And so, we were quickly in a circumstance in which there was the need to sample (via simulations or analytically) deliberative situations *and* check for the presence of order dependence in each of them. Of course, the intention behind this strategy was that perhaps we could (as we did later) report a similar result to the one from Hartmann and Rafiee Rad. However, this time we were not curious about a single model but about a large family of deliberative situations. In this approach, the notion of *opinion tree*, which is embedded in our idea of model of deliberation was a natural candidate for our sampling-and-check requirement. Naturally, the three models that we mentioned here (just before) will be properly introduced as prominent instances of models of deliberation in Chapter 1.

Sixth. The term "anchoring" is a well known one. For this reason and because we actively use it in this work, we should be cautious, so that any potential misunderstanding is prevented. In the literature (see Mussweiler et al. 2004 *and* Furnham and Boo 2011), the term "anchoring" is associated with an extensive family of biases. Also, each of these biases occurs in decision-making processes, which sometimes take place at individual level but often in group-scenarios too. The distinctive characteristic of the anchoring-family is that the judgement of decision makers is affected by an event (or "anchor") that is totally irrelevant to the situation in question (this is not good!). In the context of group deliberations, the identification of the order of the speakers with a potential anchor-event has been made explicit in Hartmann and Rafiee Rad 2020. Additionally, the anchoring effect (in non-deliberative situations) has been studied in both theoretical frameworks and empirical scenarios. This circumstance is the main reason for us to adopt the following convention here. In the context of models of deliberation, whenever we need to refer to the *potential impact* of the order of speakers on the outcome of deliberations, we have tried to use the term "*order dependence*" only. Moreover, we have reserved the term *anchoring* for "important names" exclusively (e.g., names of problems, results, sections, or the title of the dissertation). This distinction seeks to let some distance **between** the terminology used in the study of "order dependence" at a theoretical level on models **and** the names of scenarios that might be potentially relevant in an empirical study as well.

Seventh. Unfortunately, in our literature search we have not been able to pinpoint any standard source of empirical evidence that reported a result on the existence of order dependence in deliberative situations (as we present them here). Naturally, it would be great if our work promoted an interest that translated into new empirical research projects. Our hope for the existence of empirical evidence is easy to understand; it usually challenges theoretical results and intuitions in a unique way. Here too, we could have profited from some empirical guidance at the model-design stage (i.e., Chapter 1).

Despite the negative spirit of the previous paragraph, there is some literature on management science, which partially overlaps with our interests. That is, Hartmann and Rafiee Rad (see Hartmann and Rafiee Rad 2020, p. 4) have reported suggestions of a possible occurrence of (deliberative) anchoring in the particular context of decision-making boards of companies (for original sources, see Malhotra et al. 2015, Tuschke et al. 2014, Zhu 2013, and Bazerman 2002). These instances are reported as anecdotal evidence (of anchoring in deliberations), and we include them here for the sake of completeness.

Besides decision-making boards, a second interesting environment in which there seems to be a potential for anchoring is the one in which each member of a *group of experts* is asked to predict/estimate the value of a certain unknown variable that plays a role in a particular real-life situation. In these circumstances, an usual and "innocuous" assumption is that the output-predictions are formulated based on background information that belongs in the area of expertise of our specialists. Surprisingly, there is a number of issues that might arise in this simple scenario (Burgman 2016 contains a detailed inventory). In particular, anchoring is described as one among several psychological biases that might affect this kind of processes, and actual descriptions of experiments are provided in Burgman 2016, p. 60.

In our (order dependence) deliberative context, these two groups of instances are of value (even if the described settings are either not conclusive *or* not directly related to anchoring in deliberative situations) because they show the scenarios that one might expect from an empirical study on anchoring in deliberations: everyday life situations, like those from company boards **and** very specific cases, like the one in which a group of experts decides on the value of a key parameter.

Eighth. We believe that this dissertation is the first work **that** reports the result described in the answer to the *third* question above. Moreover, our general notion of model of deliberation **and** the form of approaching the study of deliberative situations (via *opinion tree* sampling) are interesting contributions to the analysis of group deliberation too.

We also believe that what is exposed here have immediate theoretical and practical relevance. On the theoretical front, our results call for an update of the known concerns regarding the limitations of collective deliberation, which is an essential component of deliberative democracy (see Estlund 2012, p. 397-406). On the practical front, our results provide enough ground, so that one could justify the need for an empirical study on anchoring in deliberations. That is, previously, we have identified the environments of deliberative negotiation, deliberative assessment of catastrophic risks, and more importantly, instances of deliberative situations within frameworks of representative democracies as potential decision-making processes (among others) that benefit from group deliberations. Naturally, there is a certain tension between this information and our results. Consequently, we consider that an empirical study that could bring clarity on this matter is needed, and it could be triggered as a reaction to the results of these dissertation.

Three pillars of Social Choice Theory

Usually, a link between deliberations and *Social Choice Theory* comes from the element of disagreement. What could we do with a deliberation that did not end in a consensus after a monumental deliberative effort? Naturally, this situation is particularly pressing when a final decision is imperative. Social Choice Theory offers a normative answer to this question, and it comes with an emphasis on the terms "voting system" and "aggregation". Unsurprisingly, both terms play a central role in other standard inquiries of the field too (see List 2022): how can a group of individuals pick one among several options? What are the strengths and weaknesses of an interesting voting system? When is a voting system fair? How to aggregate preferences of individuals in a coherent way so that the outcome represents the group?

Consequently, a short answer to the original question would be: if a deliberation does not end in consensus, there is always the opportunity to set up a voting system for making a final decision. But, which one?

Historically, the domain of research on voting systems has been bounded and supported by three prominent results (see List 2022): Condorcet's Paradox, Condorcet's jury theorem, and Arrow's Theorem. Regarding the feasibility of an aggregation of preferences via voting, these results can be taken as "negative", "positive", and "negative" respectively.

The Condorcet's Paradox can be best illustrated with an example (the original one is in Nisan 2007, p. 211). Because of a certain reason, a society needs to select one among three options (e.g., a, b, c). This particular society is very small – just three members, and their preferences for the options are as follows:

$$a \succ_1 b \succ_1 c$$

$$b \succ_2 c \succ_2 a$$

$$c \succ_3 a \succ_3 b$$

In this case, the preference $x \succ_i y$ means that the individual/voter i prefers the option x to the option y. Next, assuming that the individuals did vote according to their preferences, a quick check of the results leads us to the following outcome: two out of three voters (i.e., a majority) would prefer option "a" over option "b". Similarly, two out of three voters preferred "b" to "c", and two out of three preferred "c" to "a". Consequently, a natural interpretation of this result is that whenever a group needs to select one among several options, the intuitive majority rule should not be assumed to be a sound mechanism.

Even if the Condorcet's Paradox speaks against the idea of a majority vote as a general solution for the problem of aggregation of preferences in a group, it does not speak against the existence of a more "complicated" voting system that avoided the essence of the problem presented above. This idea could seem to be further encouraged by the Condorcet's jury theorem (see Weisstein 2022), which states that given a group of voters (i.e., a "jury"), independently choosing by majority vote between a correct outcome with probability $0 \leq p \leq 1$ and an incorrect one with probability $(1 - p)$ leads to:

1. If $p > 1/2$ (so that each voter is more likely to vote correctly that incorrectly), adding more voters increases the probability that the majority chooses correctly. Also, the probability of a correct decision approaches 1 as the number of voters increases.

2. If $p < 1/2$ (so that each voter is more likely to vote incorrectly than correctly), adding more voters decreases the probability that the majority chooses correctly. In addition, the probability of a correct decision is maximized for a jury of size one.

Naturally, an important message in this theorem is that there are some environments in which the use of the majority rule provides good results indeed. In this case, we needed a setting where the options are binary (i.e., jury is "correct" or "incorrect"), and most importantly, a large jury where each member brings more information than the act of flipping a coin.

The second negative result is Arrow's theorem (see Nisan 2007, p. 213). In essence, it extends the "bad news" of the Condorcet's Paradox beyond the scope of the majority rule to a larger class of preference aggregation methods (i.e., welfare functions). In this case, the strategy is to show that over a set of more than two options and two or more voters, a welfare function that satisfies some seemingly plausible axioms does not exist. In general, obtaining impossibility results (like the Arrow's one) is still an active research topic in Social Choice Theory, and showing that "good" aggregations methods are achievable by giving up as few as possible of some "plausible axioms" is a popular subject too (see List 2022).

In our context, voting systems will not play a role in this dissertation, but its reference is suitable and needed to precisely narrow the scope of our work: the ideas presented here are only concerned with the deliberative part of a decision making process, and any further choice that might be taken with the application of a satisfactory voting procedure falls out of the scope of our results. Likewise, we do not analyze here the impact of a previous (potentially biased) deliberative situation on a posterior voting procedure. The main reason to follow this approach is that *deliberation* as a subject is a very basic element in standard decision making processes, and so far it has not kept much exclusive attention in order to understand the strengths or weakness of its structure. Naturally, we recognize that the relation between voting systems and (biased) deliberative situations might be an interesting subject to look at in a future work.

Overview

In Chapter 1, a description of an ideal model of deliberation is presented. In general terms, here we defend that the dynamics of a model of deliberation can be represented by a finite set of tree-like structures (*opinion trees*), each of them tagged with probabilities on the branches and opinions on the vertices. Consequently, in the next chapters we use this type of structures for the study of order dependence in models of deliberation.

In Chapter 2, we present two results. **The first**, if we inspect a family of plausible *opinion trees*, it leads to the belief that the opinion of the first speaker has an overall better chance of prevailing than any other opinion of the group. *Opinion trees* in this family share the same initial conditions: a group of speakers is about to enter a debate (each speaker with a given presenting position and an opinion on the topic of debate). We will also present two factors that have an impact on the previous phenomenon: the number of opinions (admitted in the deliberation) **and** the parity of *the number of rounds times the number of speakers*. **Second**, in the same scenario as before, if instead of a family of *opinion trees*, a single one was picked (uniformly at random) and then inspected, the probability that the first speaker had better odds (of winning the debate) than any other individual speaker converges to one when the number of acceptable opinions goes to infinity.

We interpret these results to be, first, an informative overview of what happens in a family of *opinion trees* with respect to the order of the speakers. Second, an assurance telling that allowing many opinions will turn that general overview into the local standard case. That is, if one allows for many opinions, with high probability: in a uniformly selected opinion tree, the first speaker will have advantage.

In Chapter 3, we turn our attention to a more pragmatic scenario. That is, we (assume that we) are given a specific instance of a model of deliberation, and we need to answer the question: *to which extent does this instance present order dependence?* Naturally, the results in the previous chapter is what motivates this question. The reason is simple, if we know that when we inspect a family of *opinion trees*, we find that the first speaker has advantage, the next interesting question is: *how do I know whether a model that I meticulously developed presents anchoring as well?* In contrast to the results of the previous chapter (which are analytical), this one falls into the category of simulation results. In more detail, we present a general approach (that uses simulations) for estimating the degree of order dependence (and anchoring) that a model of deliberation presents.

Chapter 4 is essentially exploratory. That is, after presenting some negative results in previous chapters, in this one we investigate four natural ideas that (potentially) might help to decrease the impact of the first speaker's opinion on the outcome of a deliberation. As it might be expected, these alternatives will be described as modifications of the initial notion of model presented in Chapter 1.

Finally, in the last chapter, we summarize the main ideas of the dissertation. However, this time we try to highlight the positive side of our results: to gain knowledge about the impact of the order of speakers on deliberations, which hopefully allowed us to produce better models and deliberative situations.

Chapter 1

Models of deliberation

In this chapter we introduce a representation of what we understand as a model of deliberation. The intention is not to present an instance of a model of deliberation, but a formalism that captures a large family of interesting models instead. Consequently, in the first section we propose an intuitive description of what a model of deliberation is. Based on this, we identify the main individual components of a model, and present a formal description for each of them, so that when together, they constitute a description of model.

1.1 Models and representations

A model of deliberation M is a model that represents the situation in which n speakers deliberate on a topic during K rounds of debate (at most). The speakers enter the deliberation with an initial opinion each, and what follows respects this dynamics: the first speaker presents her opinion, then all the others update theirs. Next, the second speaker presents his opinion, and all the others update theirs. When the last speaker spoke and all the others updated, the previous presenting/updating process is repeated as long as no consensus is reached for a maximum of $K - 1$ rounds. Consensus can always emerge after a presentation, and if that was the case, it ends the debate.

What makes a model of deliberation unique is: first, the **number of speakers**. Second, the set of **possible opinions** that speakers can produce. Third, the particular form in which speakers **update their opinions**. Fourth, the mechanism for deciding whether or not the opinions are in **consensus**. Fifth, the **number of rounds** of a debate.

Next, we present a formal structure that aims to capture the elements and dynamics of plausible scenarios of deliberation. Starting with the elements: we accept any finite representation as a plausible description of an opinion. This is a gentle constraint; ultimately, we can always "define" names for an infinite opinion (the word phi is perhaps the best example). Formally:

Definition 1 (Set of opinions). Any set with a finite representation for its elements can be regarded as a plausible set of opinions. Consequently, each of its elements will be understood as an individual opinion itself.

The next essential step is to describe the **transitions between group of opinions** in a model of deliberation. We take this step by introducing what an **opinion tree** is. Graphically, an opinion tree is a structure similar to the one presented in Figure 1.1. In this particular case we have two opinions only (represented by "0" and "1"), three speakers (this is the reason why vertices are labeled with vectors of opinions with 3 components), and the debate has one single round (each speaker speaks once at most). In the figure, the position of the speaker is highlighted with a "*". An important detail is that edges between configurations of opinions are annotated with probabilities (these are **constant values** in $[0, 1]$); the intuitive meaning is that a transition between two vertices occurs with the specified probability. Regarding probabilities, there is a constraint that was not explicitly specified in the example, but it must be in place: for each vertex, its outgoing probabilities must add up to 1 (this is the standard "unit measure" requirement). Finally, in our example, the bold vertex signals the initial opinions of the speakers, which is the label of the root of the opinion tree. Now, the general notion depicted in the example can be formalized as follows:

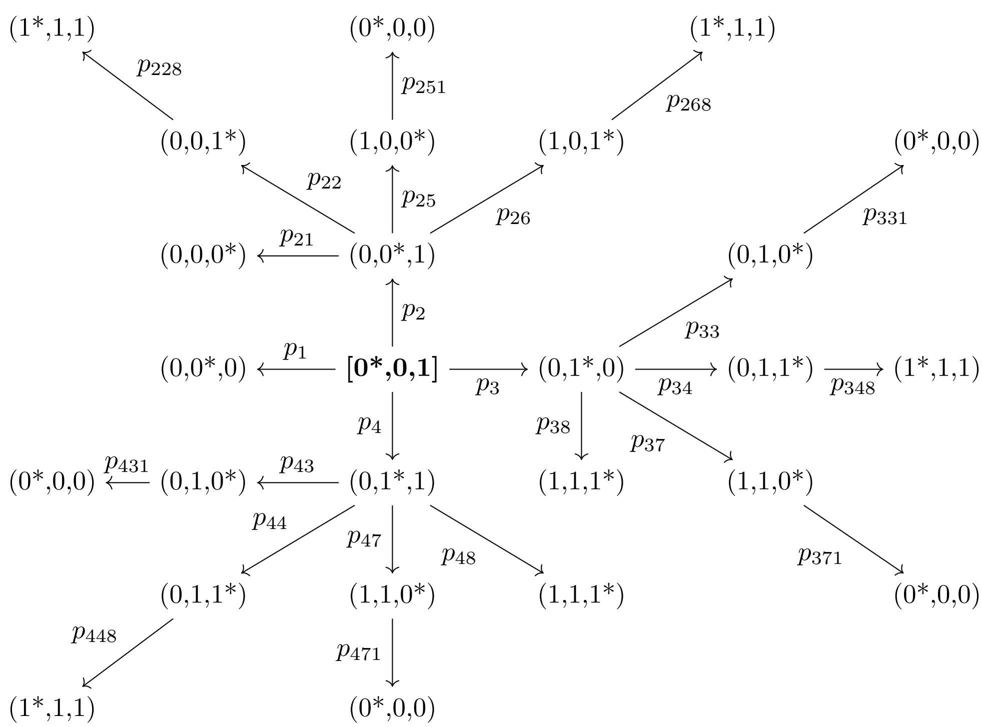

Figure 1.1: Opinion tree with three speakers, one round and two opinions.

Definition 2 (Opinion tree)**.** Given a set O of opinions and two natural numbers $K > 0$ and $n > 1$ (K stands for the number of rounds, and n for the number of speakers), we define an opinion tree as a probabilistic and labeled tree T such that: (i) O^n is the set of admissible labels. (ii) internal vertices have exactly O^{n-1} children, (iii) siblings vertices do not share labels, (iv) each vertex at a distance $d > 0$ from the root, has a label that at position $(d-1)(\textbf{ mod } n)$ has the same opinion that it's parent's label at the same position. (v) There is no leaf with a path to the root that is larger than $K \cdot n$.

Conditions (i), (ii) and (iii) describe the requirement that vertices are labeled with vectors of opinions (and only with them). Condition (iv) encodes an expected requirement: after a speaker presents, her opinion remains the same. Finally, condition (v) expresses the constraint on the number of rounds (at most K).

Two details regarding notation: first, if we had variables instead of constant values for the edge-probabilities of an opinion tree, we use the name **opinion tree structure** to highlight that this is the case. Second, because of space reasons, (quite often) we use the term O instead of the correct one $|O|$ for describing the cardinality of a set of opinions O. Naturally, we take care that there is no ambiguity when we do this.

Back again, in the example described before, one can notice that all leaves in the tree follow the same pattern: the opinion of all speakers is the same. The reason for this harmony is that in this particular instance of debate, the consensus function dictates that consensus means total agreement. In general, this does not need to be the case. A broader notion of consensus can be stated as follows:

Definition 3 (Consensus function)**.** Given a set O of opinions and a natural number $n > 1$, a consensus function $C : O^n \to \{0, 1\}$ is a function that maps a vector of opinions onto a binary value.

Intuitively, the idea behind this definition is that with the assistance of a consensus function, it is possible to classify any opinion of a group of speakers into "agreement" or "disagreement". If needed, we interpret 1 as agreement and 0 as disagreement. Also, note that the term O^n takes the usual interpretation: it represents the vectors of n components, where each component takes values from the elements of the set of opinions O.

Next, we coordinate the previous notions of *set of opinions*, *opinion tree*, and *consensus function* for describing what a model of deliberation is. This is an important step because in what follows the promise is: problems concerning the order of the speakers in a model of deliberation can be studied on the following kind of structure:

Definition 4 (Model of deliberation)**.** A model of deliberation is a tuple $M = (O, K, n, U, C)$ that consists of:

- a set of opinions O,
- a definite number of rounds K,

- a definite number of speakers n,
- a finite set of opinion trees U, with no repeated root-labels.
- a consensus function $C : O^n \to \{0, 1\}$,
- A constraint: consensus vertices have no children. That is, in an opinion tree $T \in U$, a vertex with label x is a leaf **iff** $C(x) = 1$ or its distance to the root is equal to $K.n$.

The new elements are a requirement of a finite set U of opinion trees, and a constraint. First, that U is a finite set with no repeated root-labels indicates that any pair of opinion trees in U describes debates that started with different initial group opinions. Second, the constraint narrows the definition of a leaf in opinion trees by preventing that consensus vertices had children (i.e., when consensus is reached the debate stops).

This definition will be used as our ideal of a model of deliberation. Accordingly, we use it for studying the relevance of the order of the speakers to the final outcome of a debate. However, this subject will not be treated directly; it will be first unfolded into two kind of questions instead: questions about individual models (in Chapter 3) and questions about groups of opinion trees (in Chapter 2). But, before we get into more details on this, we first contrast our new representation of a model of deliberation with some examples from the literature.

Example 1. First, we consider any deterministic model $M_{det} = (O, K, n, U, C)$. When we say that M_{det} is deterministic, it means that: in every tree of U there is a unique path where no edge-probability is equal to zero. Additionally, in this path all edge-probability are equal to one. That is, during deliberations, M_{det} asserts with certainty the transitions between two groups of opinions. Naturally, particular deterministic models can be obtained by instantiating the parameters (O, K, n, U, C).

As the previous example suggests, by identifying different mechanisms for opinion-update, one can produce different models of deliberation (i.e., by producing different sets U). Under the deterministic classification, **two prominent instances** in the literature are the Lehrer-Wagner (see Lehrer and Wagner 1981) and the Hegselmann-Krause (see Hegselmann and Krause 2002) models. **In the first case**, the updating mechanism is linear. That is, updates are done via a matrix multiplication with the general form $V^{i+1} = WV^i = W^iV_0$. This expression can be read as: after the $(i + 1)$ round of debate, the opinions of the speakers (V) are equal to the multiplication of a matrix W by the opinions of the speakers after the previous round. The matrix $W = w_{mk}$ can be interpreted as an encoding of opinion significance among speakers. That is, w_{mk} expresses how significant is the opinion of speaker k for speaker m. An interesting consequence that results from this model is that: for a large family of matrices W (independently of the initial opinion), the debate converge to $\lim_{i \to \infty} V^i$. **In the second case**, the mechanism for updating opinions is uncomplicated as well. In every round, a speaker updates her opinion with the average of all opinions that are similar to hers. The notion of "similar" is controlled with a parameter, usually named ϵ, and interpreted as a confidence threshold (an individual feature of speakers). Some known implications of this updating style are as follows: first, a stable pattern of opinions is obtained after a finite time. Second, the shape of the concluding patterns heavily depends upon the ϵ profile of speakers. Consequently, by tuning ϵ, different patterns of plurality, polarization or total agreement can be obtained.

Could we describe these models in a similar way to that of Definition 4? Well, they are deterministic. Accordingly, what was said in Example 1 applies to them as well. However, the significance of these models was mainly due to "on the limit" results (i.e. the length of deliberations was not bounded). That is, if we would like to describe them in the style of Definition 4, we should allow (on the definition side) for infinite trees in U (in the sense of an infinite K). A second difficulty would be that both models describe the update of opinions in a synchronous way. That is, all opinions are updated at the same time (and updates are broadcast instantly). In our case, it is expected that each round of debate contains as many opinion-update possibilities as speakers exist (unless consensus is reached). This turns out to be a fundamental difference that we can not overcome without making too many concessions on our (above) ideal of debate. We come back again to this point in a moment when we explicitly set the scope of the results presented in the next chapters. Next, a second example, of a non-deterministic model this time.

Example 2. In the HRR model of deliberation (see Hartmann and Rafiee Rad 2020), the dynamics of a debate is described as follows: deliberations take place in rounds, and speakers present their opinions sequentially (as required by Definition 4). More important, speakers exhibit and make use of two key features (unnoticeable for themselves, and treated as parameters in the model): a first order reliability and a second order reliability. These attributes are used for explaining how is that speakers update their opinions. Consequently, the first order reliability is interpreted as the ability to take right decisions (i.e., select the right among many possible options), and the second order reliability is interpreted as the capability to estimate the first order reliability of other speakers. Both attributes change their values in a precise (algebraic) and justified way while a debate unfolds. In this environment, some interesting implications are as follows: first, **if the speakers are epistemic peers** (this is expressible in terms of their first order reliability), then the deliberation process converge to a consensus. Equally important, in the way to this consensus, the opinion of the i^{th} speaker receives a higher weight than the one of any other subsequent speaker. This circumstance is interpreted as a new instance of the prominent anchoring effect. **If the speakers are not epistemic peers**, it has been demonstrated (via simulations) that for many plausible scenarios the same anchoring effect is exhibited.

How can we contrast the HRR model with Definition 4? Regarding the structure of debates both are alike: debates happen in rounds, speakers present their mind sequentially, and the numbers of rounds is bounded (i.e., no infinite

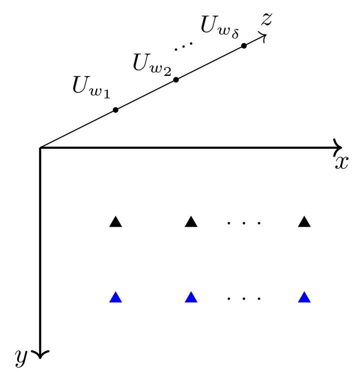

Figure 1.2: Graphic description of a family of models of deliberation.

trees). However, there is an interesting difference regarding consensus: Definition 4 allows for debates that ended without consensus, and in the HRR model, an average of opinions is taken in that case (i.e., if a maximum number of rounds is reached without consensus). Could we have this behavior on the definition side? Yes, it is possible. In order to do that, one should extend the domain of the consensus function, so that it took into account not only the plausible labels (i.e., group-opinions), but the length of deliberations as well. Naturally, labels in leaves (of opinion trees) should then take the value of the average of labels in their parent vertex (note, labels' values needed to be numerical in this case).

So far, we have examined individual instances of models, next we inspect a group view instead (taking Definition 4 as the notion of singleton). As the previous examples indicated, the set U of trees can be seen as the main component of a model of deliberation. For a typical model, U might be large, but if (O, K, n, C) are fixed, U must be finite. In this case (i.e., (O, K, n, C) fixed), we can informally picture a family of models of deliberations to be something close to what Figure 1.2 shows (triangles stand for opinion trees): each "slice" of the xy-plane contains a particular U, and in the z-direction each *opinion tree structure* takes different probability assignments on the edges. Additionally, we could also give some organization to each individual U by accommodating under the same y-coordinate those trees with root-labels that are a permutation of each other (represented by equal colors in our example). For instance, if we consider the tuple of opinions $[0, 0, 1]$ as a root-label (see Figure 1.1), the opinion trees with root-labels $[0, 0, 1], [1, 0, 0]$ and $[0, 1, 0]$ will share the same y-coordinate. The profit of keeping this view will become apparent in Chapter 3.

As mentioned earlier, in the next chapters there will be **two different kind of problems** concerning the order of speakers in models of deliberation. **The first one** (discussed in Chapter 2) is a group kind of problem. In this case, the situation is that a group of speakers is about to enter a debate (each speaker with a given presenting position and an opinion on the topic of debate), and we would like to know whether the opinion of the first speaker has an overall better chance of prevailing than other opinions of the group. We model this problematic by investigating (on the corresponding *opinion tree structure*) whether the expectation that the opinion of the first speaker prevails is larger than the expectation that any other individual opinion did it. In terms of Figure 1.2, this would be: for a particular *opinion tree structure* (that agrees with the situation to model), we sample its corresponding opinion trees over the z-direction (in a uniform and independent way), and keep track of **the average** of the **difference** of probability of success **between** the opinion of the first speaker and any other opinion. As revealed before, the answer to this problem is that the first speaker has advantage (i.e., the average is positive). Moreover, we will also show that the same holds (with high probability) if the set O of opinions is large, and we did not sample over the z-direction, but just selected (uniformly) a single opinion tree instead.

The second kind of problem (discussed in Chapter 3) is about single models. In this case, we are given a particular model (i.e., a "slice" in the xy-plane) as a **black box**: we know (O, K, n, C). Additionally, we have access to the state of opinions at any moment during a deliberation (but we do not have access to the probabilities of trees in U), and we would like to know to which extent there was order dependence in this model. Naturally, in the corresponding chapter we properly explain the motivation to consider this problem, and describe our approach to it.

Discussion/Conclusion. In order to complete this chapter, in the next paragraphs we highlight the main ideas that were presented earlier. We also review important points that were briefly mentioned (with the intention of keeping the description of models as clear as possible), but still might need further clarification. For this we use a simple dialogue style.

Why is the differentiation between opinion trees and opinion tree structures needed? An opinion tree is used for describing a very particular instance of a debate. In this case, the probability of transitions (between two group-opinions) are supposed to exist as constant values. Opinion tree structures highlight the idea that the parameters (O, K, n, C) have fixed values, and the probability values of transitions between group-opinions are kept uncertain and described as random variables.

Is it possible to get access to the edge-probabilities of opinion trees? If we are given a model of deliberation described in natural language (or even by mathematical rules), the answer is no (in general): a basic reason is that when dealing with probabilities we can not guarantee certainty (i.e., via simulations). However, we can estimate those values, most of the time with some practical difficulties (i.e., U might be a large set, and the trees in U might be large as well). Moreover, for simple models this problem can be tractable, and depending on the regularity of the rules, probability transitions might be easy to derive (analytically). We will come back to the general question, and discuss it again in detail in Chapter 3 .

What is a family of models exactly? For our purposes, we will not need to define this in a formal way. However, (informally) Figure 1.2 captures the idea in a genuine way. That is, it is a collection of models that can be obtained by fixing (O, K, n, C) and allowing all possible scenarios for the edge-probabilities of opinion tree structures.

Are the results of the next chapters presented in full generality? No. That is, they are not shown for every tuple (O, K, n, C). We will assume everywhere that $C(x)$ is a **total agreement consensus function**. Specifically, $C(x) = 1$ **iff** all the opinions in the label x are the same (we call x a **total consensus label**). Also, in the next chapter, Theorem 5 needs a large set O of opinions to be of a practical use. However, these restrictions should not be understood as strong constraints: first, in most cases agreement means exactly that all speakers share the same mind (i.e., $C(x) = 1$). Second, the assumption of a large O is a natural one (in particular in debates where diversity of opinions is allowed). Moreover, we will demonstrate (via simulations) that for cases with small sizes of O, a similar pattern (to the one with a large O) of order dependence is obtained.

Chapter 2

Order Dependence and Anchoring

In this chapter we address two core questions of the dissertation: the *overall anchoring problem* and the *probability of anchoring problem*. The solutions of these problems support the idea that **in a situation in which a group of speakers is about to enter a debate** (each speaker with a given presenting position and an opinion on the topic of debate), **the opinion of the first speaker has an overall better chance of prevailing than other opinions of the group**. Moreover, **the same is true (with high probability) if: in a scenario with many opinions, a single opinion tree was selected (uniformly)**.

In more detail, the *overall anchoring problem* (Section 2.2) asks whether in an uniform sampling of plausible opinion trees, the opinion of the first speaker prevails (on average) over any other opinion. Moreover, the *probability of anchoring problem* (Section 2.3) asks for the **probability that** the opinion of the first speaker prevails over any other **in** a single opinion-tree-pick that was sampled uniformly. Naturally, before we can turn our attention to these exciting problems, we need first to formalize the intuitive notion of "probability that a certain outcome occurs on a single opinion tree" (Section 2.1). The reason to do so: whenever we said "prevail" before, the probabilistic nature of the statement was intentionally concealed, and it needs to be made precise.

2.1 Probabilities in opinion trees

In probability theory, the notion of probability space is at the core of nearly every inquiry: if we need to describe what we mean by "the probability that an event occurs", (more often than not) we need to present a probability space that models our situation (see the Probability Theory section of Basic Notions). In this section we present a probability space that is associated to a given opinion tree. By doing so we obtain a safe interpretation for the statement "the probability that an outcome A occurs in a given opinion tree is P". But, let's see first an example that motivates the required definitions.

Example 3. We consider the example introduced in Figure 1.1 where the setting was: three speakers, two opinions, a single round of deliberation, and initial opinions **[0*,0,1]** . The challenge is to describe the probabilities $P_{(0,0,0)}$ and $P_{(1,1,1)}$ corresponding to the events: "$(0,0,0)$ is the outcome of the deliberation", and "$(1,1,1)$ is the outcome of the deliberation".

Intuitively, for describing $P_{(0,0,0)}$ and $P_{(1,1,1)}$, the procedure is to take the ratio of the outcomes we are interested in to all possible consensus outcomes. If we expand the previous idea, it takes the following form:

$$P_{(0,0,0)} = \frac{S(0,0,0)}{S(1,1,1)+S(0,0,0)}$$

$$P_{(1,1,1)} = \frac{S(1,1,1)}{S(1,1,1)+S(0,0,0)}$$

$$S(0,0,0) = \hat{P}_{251} + \hat{P}_{331} + \hat{P}_{371} + \hat{P}_{471} + \hat{P}_{431} + \hat{P}_{1} + \hat{P}_{21}$$
$$S(1,1,1) = \hat{P}_{268} + \hat{P}_{348} + \hat{P}_{38} + \hat{P}_{48} + \hat{P}_{448} + \hat{P}_{228}$$

$$\begin{aligned}
\hat{P}_{251} &= p_2 \cdot p_{25} \cdot p_{251} \\
\hat{P}_{331} &= p_3 \cdot p_{33} \cdot p_{331} \\
\hat{P}_{371} &= p_3 \cdot p_{37} \cdot p_{371} \\
\hat{P}_{471} &= p_4 \cdot p_{47} \cdot p_{471} \\
\hat{P}_{431} &= p_4 \cdot p_{43} \cdot p_{431} \\
\hat{P}_{21} &= p_2 \cdot p_{21} \\
\hat{P}_{1} &= p_1
\end{aligned}
\qquad
\begin{aligned}
\hat{P}_{268} &= p_2 \cdot p_{26} \cdot p_{268} \\
\hat{P}_{348} &= p_3 \cdot p_{34} \cdot p_{348} \\
\hat{P}_{448} &= p_4 \cdot p_{44} \cdot p_{448} \\
\hat{P}_{228} &= p_2 \cdot p_{22} \cdot p_{228} \\
\hat{P}_{38} &= p_3 \cdot p_{38} \\
\hat{P}_{48} &= p_4 \cdot p_{48}
\end{aligned}$$

The next definition simply combines the terms $S(0,0,0)$ and $S(1,1,1)$ under a single name. Note that in an opinion tree, the weight of a leaf is defined to be the multiplication of the probabilities on the path from the leaf to the root of the tree (see the Graph Theory section of Basic Notions).

Definition 5 (Outcome function)**.** The outcome function $f : O^n \to \mathbb{R}$ of an opinion tree T is defined as follows: given a label x, the value $f(x)$ is the sum of the weights of leaves with label equal to x in T.

Similarly, a single name for $P(0,0,0)$ and $P(1,1,1)$ is defined as follows. Note, in our particular example the constant α is simply $S(1,1,1)+S(0,0,0)$, and the consensus function requires total agreement.

Definition 6 (Normalized outcome function)**.** Given an opinion tree T, its outcome function $f : O^n \to \mathbb{R}$ and a consensus function $C : O^n \to \{0,1\}$, the normalized outcome function g is defined as follows: the domain of $g(x)$ are the labels $x \in O^n$ such that $C(x) = 1$. Given a label x, $g(x) \coloneqq \frac{1}{\alpha} f(x)$, where $\alpha \coloneqq \sum\limits_{C(y)=1} f(y)$ is a normalization constant.

Naturally, in the language of the previous definitions, we can describe what we mean by "the probability that an outcome occurs in an opinion tree". More important, after the next definition we can safely produce statements of the form: given an opinion tree, the probability that the opinion of the first speaker prevails is higher than the probability that any other opinion did it. For instance, in the previous example, that happens whenever $S((0,0,0)) > S((1,1,1))$.

Definition 7 (Associated probability space)**.** Given an opinion tree T, a consensus function C, and their normalized outcome function $g(x)$, the triple $(\Omega, \mathcal{F}, P)$ is the associated probability space defined as follows:

- $\Omega = \{x \in O^n | C(x) = 1\}$
- $\mathcal{F} = 2^\Omega$
- For $E \in \mathcal{F}$:

$$P(E) = \begin{cases} \sum\limits_{x \in E} g(x) & E \neq \emptyset \\ \\ 0 & E = \emptyset \end{cases}$$

.

We can appreciate that the triple $(\Omega, \mathcal{F}, P)$ is a probability space: first, $\mathcal{F}$ is a σ-algebra (it is the power set of a finite set). Second, $P(\Omega) = 1$ follows from the definition of $g(x)$. Third, if $A_i \in \mathcal{F}$ is a countable sequence of disjoint sets (because $\mathcal{F}$ is finite, this sequence needs to be finite indeed), then (based on the assumption that $A_i's$ are disjoint) $P(\cup_i A_i) = \sum_i P(A_i)$. Note, for the entire definition to make sense, there should be at least one leaf x in T with $C(x) = 1$. Otherwise, $\alpha = 0$ and $g(x)$ is not well defined.

Back to opinion trees, we can now continue as planned: so far we have considered opinion trees with literal probability values assigned to the edges between vertices (that is, literal numbers—even if in the previous examples those numbers were represented symbolically by names). In the next section we introduce the *overall anchoring problem*; this problem considers the case in which the edge-probabilities of an opinion tree are random variables (i.e., we explore an *opinion tree structure*). By taking probabilities in this way, we can sample individual trees (uniformly), and study **the difference between** the probability of success of the first speaker **and** the probability of success of any other speaker. In terms of our previous example, a first step of this study would be to inspect the term $\lim\limits_{m\to\infty} \frac{\sum_{m>0} S_m(0,0,0) - S_m(1,1,1)}{m}$ (where $S_m(0,0,0)$ and $S_m(1,1,1)$ is the evaluation of $S(0,0,0)$ and $S(1,1,1)$ for the sample m).

Naturally, one can notice that this expression is (with probability one) nothing else but the expectation of the random variable $D = S(0,0,0) - S(1,1,1)$. That is, the next section will be dedicated to the problem of calculating $\mathbb{E}(D)$ in an *opinion tree structure* with arbitrary parameters (O, K, n).

2.2 Overall anchoring

In this section there will be **three highlight moments**: first, the description of the random variable that we are going to consider in the *Overall anchoring problem*. Second, the definition of the problem itself and the presentation of its solution. Third, the discussion of the solution together with some interesting implications.

As a motivation, let us consider the *opinion tree structure* from Example 3. An interesting (but still informal) anchoring related question would be: if we sample (uniformly) the values of the edge-probabilities, and for each sample m we obtain $D_m = S_m(0,0,0) - S_m(1,1,1)$, will the average of the D_m values be positive or negative (for a large enough m)?

From an anchoring point of view, the previous question is attractive because a correct answer to it would suggest how strong the dominance of the first speaker position is. That is, we will be considering different opinion trees (in our particular example, all of them had $[0*, 0, 1]$ as root label), and for each of them we would keep track of the difference of "strength" between the opinion of the first speaker and any other opinion. Note, the introduced requirement that the sampling needs to be uniform just means that all well formed opinion trees are equally plausible.

From an operational point of view, the previous question is interesting because if we secure a probability space that captured the previous scenario (i.e., a space where we can sample opinion trees uniformly), the anchoring related question can easily be translated (via the Strong Law of Large Number) into a formal probability question. That is, we were just be looking for the expectation $\mathbb{E}(D)$ of the random variable $D = S(0,0,0) - S(1,1,1)$. Further, if we describe the quantity D precisely as a random variable (so far our scope was the Example 3 only) the problem of finding $\mathbb{E}(D)$ is precisely what we will define as the *Overall anchoring problem*.

For the general case, the claim to support is: D is defined as a combination of additions, multiplications and a single subtraction of the probabilities in an *opinion tree structure*, and because these probabilities are random variables so will be D. Next, let us expand on the idea that the edge-probabilities are random variables (a confident reader who is sure that D is indeed a random variable can dodge the next three paragraphs and jump directly to our next topic).

Over the edge-probabilities there are three simple but important constrains: first, probabilities under the same parent should add up to one. Second, their values are sampled uniformly. Third, any two groups of probabilities under two different parents should be independent. The origin of these constraints is the situation that we are modeling. That is, the first one captures the fact that for each opinion vertex the outgoing branches cover all possible future opinions. The second expresses that all well formed opinion trees are equally likely to be considered. The third one describes that when we sample opinion trees (by sampling their edge-probabilities), the selection of probability values for edges under different vertices do not affect one each other. But, now the question is: are those edge-probabilities random variables, and if yes how does the probability space of these random variables look like?

The answer can be explained as follows: from the theory of probabilities we know that given a multivariate distribution function F (with arity k), there exists a random vector (with the same arity) that distributes according to F (see Billingsley 2012, p. 276-277 and p. 199). But, do we have a multivariate distribution function for the edge-probabilities? The answer is yes. From the first two constraints above, we know that probability values under the same parent follow a Dirichlet distribution (with the parameter α equal to the k-ary unit vector). Thus, probability values under a parent v form a random vector X_v with $(\mathbb{R}^{k-1}, \mathcal{R}^{k-1}, \mu)_v$ as its sample space; μ can defined in terms of F (see Billingsley 2012, p. 187), and $X_v(w) \equiv w$.

Now, we just need to combine the previous Dirichlet distributions (we have one per internal vertex v) in a way that satisfies the third constraint (the one about independence). The strategy to follow is standard for producing a finite number of independent random vectors when each of them must follow a particular distribution (Durrett 2019, p. 45). This strategy has three key ideas: first, there will be a new measurable space, and it is large (will be the product of $(\mathbb{R}^{k-1}, \mathcal{R}^{k-1})_v$). Second, the measure of the new probability space is defined as the product of the old measures (each one expressed in terms of its given distribution function). Third, the actual independent random vectors are defined as $X_i(\omega_1, \omega_2, \cdots, \omega_v) = \omega_i$ (note, w_i is a vector). Finally, we can express D in terms of the components of the $w_i's$.

Assuming that we are confident that D is a random variable (with the above mentioned requirements), next we can proceed with the second important moment of this section: the problem of calculating the expectation $\mathbb{E}(D)$ of D.

Definition 8 (Overall anchoring problem)**.** Given an *opinion tree structure* with parameters (O, K, n), the overall anchoring problem asks for the expectation of the random variable D defined as: **the difference between** the outcome function evaluated on a *total consensus label* of the opinion of the first speaker **and** the outcome function evaluated on a *total consensus label* of any other opinion.

Before we present the solution to this problem, let us notice that if we were not looking for $\mathbb{E}(D)$ but for an estimation of it instead (for a particular instance of (O, K, n)), we could proceed as follows: we create the opinion tree structure corresponding to the parameters (O, K, n), and generate a uniform Dirichlet vector of probabilities for each internal vertex (i.e., for its outgoing edge-probabilities). Then we calculate the value of D, and repeat the same experiment several times. The average of all Ds will be our estimator of $\mathbb{E}(D)$ for this particular instance of (O, K, n). This is precisely what Table 2.1 shows, estimators of $\mathbb{E}(D)$ for different instances of the parameters (O, K, n) (particulars about the generation of the data can be found in Appendix D.1 (Experiment1)). Considering these data, there are two aspects that we can expect from a general solution of the overall anchoring problem. First, $\mathbb{E}(D)$ seems to be positive. Second, with K, n fixed, $\mathbb{E}(D)$ appears to be strictly decreasing on O.

Theorem 1. *Given an opinion tree structure with parameters* (O, K, n) *and a total agreement consensus function, if we take* $p := \frac{1}{|O|^{(n-1)}}$ *and* $\ell := K \cdot n$*, the solution of the Overall anchoring problem is:*

$$\mathbb{E}(D) = \begin{cases} \frac{p+p^{\ell+1}}{p+1} & \ell \text{ is odd} \\ \\ \frac{p-p^{\ell+1}}{p+1} & \ell \text{ is even} \end{cases} \tag{2.1}$$

From this characterization of $\mathbb{E}(D)$, there are two further aspects that we could add to those already expected (from the information in Table 2.1). First, the parity of ℓ plays a role in the value of $\mathbb{E}(D)$. Second (we will use this detail later), for large values of O, the value of $\mathbb{E}(D)$ is very small, and it behaves like p. That is (see Appendix C.1 for the details),

$$\lim_{O \to \infty} \frac{\mathbb{E}(D)}{p} = 1 \tag{2.2}$$

Back to our result, the central idea for proving Theorem 1 is that: in an *opinion tree structure*, the expected values of the edge-probabilities are all equal to p. This is a direct consequence **of** the uniform Dirichlet distribution that was assumed for the edge-probabilities under each internal vertex (for details on the expectation in this distribution, see the Dirichlet distribution section of Basic Notions) **and** the independence constraints mentioned above too. Next, for making a precise use of this idea, we state three utility lemmas, and after that we are ready to prove the theorem.

Roughly speaking, the first lemma reduces the problem of calculating $\mathbb{E}(D)$ in an *opinion tree structure* to the problem of calculating D in an *opinion tree*. More precisely, it states that in a given *opinion tree structure*, the value $\mathbb{E}(D)$ is identical to the value D in the corresponding opinion tree with all edge-probabilities equal to p. As an example, let us re-visit the opinion tree in Figure 1.1. But, this time with all its edge-probabilities taking the value p as in Figure 2.1. Now, if we evaluate $S((0,0,0) - S(1,1,1))$ using their expressions from Example 3, the result is $p^3 + p - p^2$. This is indeed the same result as the one that can be obtained via Theorem 1:

$$\mathbb{E}(D) = \frac{p + p^4}{p+1}$$

$$= \frac{p + p^4 + (p^3 - p^3) + (p^2 - p^2)}{p+1}$$

[O,K,n]	Estimates of $\mathbb{E}(D)$
[2,1,2]	0.2477
[3,1,2]	0.2220
[4,1,2]	0.1904
[5,1,2]	0.1612
[6,1,2]	0.1397
[7,1,2]	0.1225
[8,1,2]	0.1088
[9,1,2]	0.0985
[10,1,2]	0.0899
[11,1,2]	0.0826
[2,2,2]	0.3111
[3,2,2]	0.2414
[4,2,2]	0.1990
[5,2,2]	0.1646
[6,2,2]	0.1438
[7,2,2]	0.1242
[8,2,2]	0.1110
[9,2,2]	0.0996
[10,2,2]	0.0909
[11,2,2]	0.0835
[2,3,2]	0.3250
[3,3,2]	0.2489
[4,3,2]	0.1994
[5,3,2]	0.1655
[6,3,2]	0.1411
[7,3,2]	0.1255
[8,3,2]	0.1116
[9,3,2]	0.0999
[2,1,3]	0.2029
[3,1,3]	0.0992
[4,1,3]	0.0584
[5,1,3]	0.0387
[6,1,3]	0.0271
[7,1,3]	0.0198
[8,1,3]	0.0153
[2,2,3]	0.1989
[3,2,3]	0.1004
[2,3,3]	0.1996
[2,4,3]	0.1978

Table 2.1: Estimates of $\mathbb{E}(D)$. The computation involved 20000 experiments for each profile.

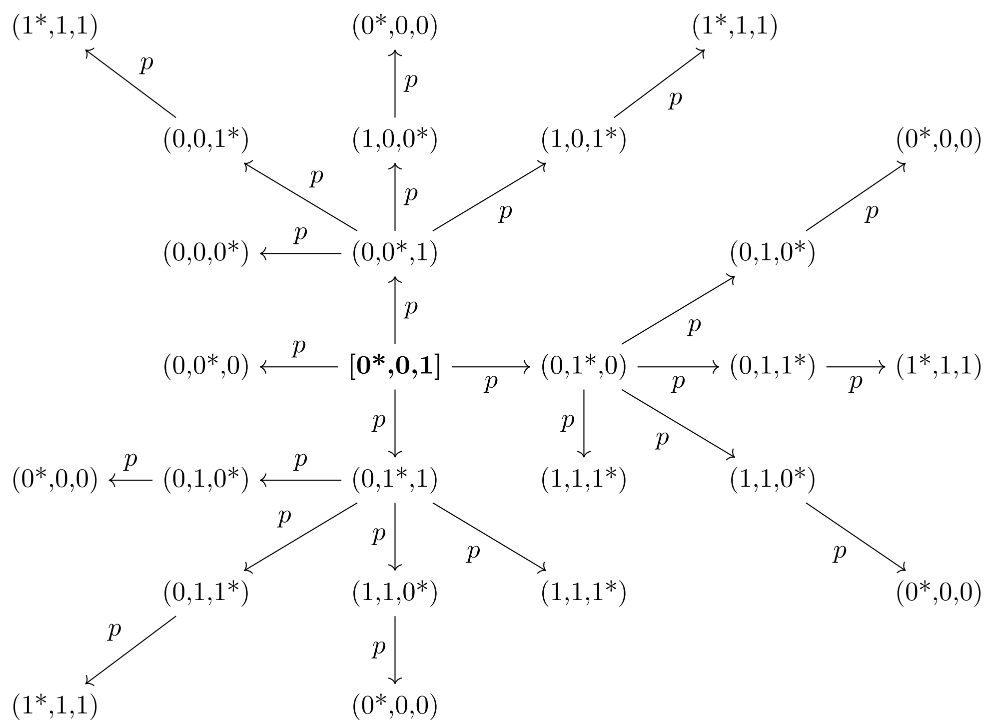

Figure 2.1: Opinion tree with three speakers, one round and two opinions. Edge probabilities are all equal to p.

$$= \frac{(p^3 + p - p^2)(p+1)}{p+1}$$

$$= p^3 + p - p^2$$

Note, it is clearly required that the sum of the values p of the edge-probabilities under the same parent add up to the unit. In this case, $4p = 1$ leads to $p = \frac{1}{4}$, which is the same value $\frac{1}{|O|^{(n-1)}} = \frac{1}{2^2}$ introduced for p in Theorem 1. The previous example can be generalized as follows:

Lemma 2. *Given an opinion tree structure with parameters (O, K, n)* ***and*** *an opinion tree with the same parameters and edge-probabilities p, the following holds: the expectation of D in the opinion tree structure is equal to the evaluation of D in the opinion tree. That is $\mathbb{E}(D) = D$.*

See Appendix B for a proof of Lemma 2.

With this lemma, we have reduced the problem of calculating $\mathbb{E}(D)$ to the one of calculating D. Now, if we need to calculate D, it will be useful **to know** the number of consensus vertices (for a given opinion) that exists (with that opinion) at a distance i from the root. But, given that a consensus vertex can only be generated from an active vertex (i.e., not in consensus), we will consider the calculation of vertices $L_i^{A,a}$ that are as follows: **A**ctive, at a distance $(\boldsymbol{i}-1)$ from the root, and the current speaker has the specific opinion "**a**". That is, we are interested in the value of $L_i^{A,a}$ because it tells how many "a" labeled consensus vertices exist at a distance i from the root. The next two lemmas will take a closer look at the properties of $L_i^{A,a}$.

Without loss of generality, let "0" be the opinion of the first speaker and "a" the opinion of any other speaker ("a"$\neq$ "0"). Further, let $\overrightarrow{0}$ and $\overrightarrow{a}$ be tuples of opinions with all the values equal to "0" and "a" respectively (naturally, their length is equal to the number of speakers). Next, following its definition, D can be calculated as:

$$D = S(\overrightarrow{0}) - S(\overrightarrow{a})$$

Moreover, these terms can be unfolded as follows:

$$\begin{aligned}S(\overrightarrow{0}) &= p \cdot L_1^{A,0} + p^2 \cdot L_2^{A,0} + p^3 \cdot L_3^{A,0} + \cdots + p^\ell \cdot L_\ell^{A,0}\\ S(\overrightarrow{a}) &= p \cdot L_1^{A,a} + p^2 \cdot L_2^{A,a} + p^3 \cdot L_3^{A,a} + \cdots + p^\ell \cdot L_\ell^{A,a}\end{aligned}$$

Which results in the following characterization of D:

$$D = p\cdot(L_1^{A,0} - L_1^{A,a}) + p^2\cdot(L_2^{A,0} - L_2^{A,a}) + p^3\cdot(L_3^{A,0} - L_3^{A,a}) + \cdots + p^\ell\cdot(L_\ell^{A,0} - L_\ell^{A,a}) \tag{2.3}$$

Then, our second lemma provides us with a simple characterization for the differences $L_i^{A,\Delta} := L_i^{A,0} - L_i^{A,a}$.

Lemma 3. $L_i^{A,\Delta} = (-1)^{i+1}$

See Appendix B for a proof of Lemma 3. **A detail regarding notation:** whenever needed in the rest of the text, we use H and Q as abbreviations of the expressions $|O|^{n-2}$ and $|O|^{n-1} - 1$ respectively.

Naturally, we can see that Lemma 3 and Eq. 2.3 are in agreement with what we know from Example 3:

$$\begin{aligned}D &= S((0,0,0) - S(1,1,1))\\ &= p^3 + p - p^2\\ &= p - p^2 + p^3\\ &= p(-1)^{1+1} + p^2(-1)^{2+1} + p^3(-1)^{3+1}\end{aligned}$$

.

Because we know that $D = \sum_{i=1}^{\ell}(-1)^{i+1}p^i$, the last lemma simply provides us with two succinct expressions for this sum (depending on the parity of ℓ).

Lemma 4. *The closed form of* $\sum_{i=1}^{\ell}(-1)^{i+1}p^i$ *is* $\frac{p+p^{\ell+1}}{p+1}$ *when* ℓ *is odd, and* $\frac{p-p^{\ell+1}}{p+1}$ *otherwise.*

See Appendix B for a proof of Lemma 4. Naturally, after this we can proceed and see Appendix B for a proof of Theorem 1 too.

With the proof of Theorem 1 completed, there are still three points that we need to revisit before we close the section. **First**, Table 2.2 complements the information presented in Table 2.1. That is, in addition to the estimates of $\mathbb{E}(D)$ for different (O, K, n) profiles, now we know the actual values of $\mathbb{E}(D)$. As one might have anticipated, pairs of values with the same profile seems to be in agreement. **Second**, an example of the use of Eq. B.5 and Eq. B.6 will be shown next (on the *opinion tree structure* in Figure 2.2); these equations had an important role in this section, and will still be present in the next one. Figure 2.2 shows an *opinion tree structure* for the case in which $(O, K, n) = (3, 1, 3)$. Because of the size of the tree, the following renaming of the vertices was needed (superscripts in the vertices of the tree indicate the position of the speaker).

$$
\begin{array}{rrr}
g_1 = [0,0,0] & g_2 = [0,0,1] & g_3 = [0,0,2] \\
g_4 = [0,1,0] & g_5 = [0,1,1] & g_6 = [0,1,2] \\
g_7 = [0,2,0] & g_8 = [0,2,1] & g_9 = [0,2,2] \\
g_{10} = [1,0,0] & g_{11} = [1,0,1] & g_{12} = [1,0,2] \\
g_{13} = [1,1,0] & g_{14} = [1,1,1] & g_{15} = [1,1,2] \\
g_{16} = [1,2,0] & g_{17} = [1,2,1] & g_{18} = [1,2,2] \\
g_{19} = [2,0,0] & g_{20} = [2,0,1] & g_{21} = [2,0,2] \\
g_{22} = [2,1,0] & g_{23} = [2,1,1] & g_{24} = [2,1,2] \\
g_{25} = [2,2,0] & g_{26} = [2,2,1] & g_{27} = [2,2,2]
\end{array}
$$

In this case, the consensus states are $\{g_1, g_{14}, g_{27}\}$. Also, $H = |O|^{n-2} = 3^{(3-2)} = 3$ **and** $Q = |O|^{n-1} - 1 = 3^{(3-1)} - 1 = 8$. As expected, the evaluations of different $L_i^{A,x}$ are in agreement with what we can see in the tree. Further, the differences $L_i^{A,\Delta}$ are in harmony with Lemma 3.

$$L_1^{A,0} = \frac{(-1)^1(3.8^1(-1)^1 + 8(-8+3-1))}{8(8+1)} = \frac{72}{72} = 1$$

$$L_2^{A,0} = \frac{(-1)^2(3.8^2(-1)^2 + 8(-8+3-1))}{8(8+1)} = \frac{144}{72} = 2$$

$$L_3^{A,0} = \frac{(-1)^3(3.8^3(-1)^3 + 8(-8+3-1))}{8(8+1)} = \frac{1584}{72} = 22$$

$$L_1^{A,1} = L_1^{A,2} = \frac{3(-1)^1((-1)^1 8^1 + 8)}{8(8+1)} = \frac{0}{72} = 0$$

$$L_2^{A,1} = L_2^{A,2} = \frac{3(-1)^2((-1)^2 8^2 + 8)}{8(8+1)} = \frac{216}{72} = 3$$

$$L_3^{A,1} = L_3^{A,2} = \frac{3(-1)^3((-1)^3 8^3 + 8)}{8(8+1)} = \frac{1512}{72} = 21$$

The **last point** to revisit is something that we mentioned before presenting Theorem 1: we observed that according to Table 2.1 "with K, n fixed, $\mathbb{E}(D)$ appears to be strictly decreasing on O". Next, we will see that because the derivative of $\mathbb{E}(D)$ wrt. to O is negative (with the exception of a single point where it is zero), this idea was indeed correct (see Appendix C.4 for details on the derivative).

$$\frac{\partial \mathbb{E}(D)}{\partial O} = \frac{(n-1)\boldsymbol{(-1 + (-(\frac{1}{O})^{(n-1)})^{K.n} + K.n(1 + (\frac{1}{O})^{(n-1)})(-(\frac{1}{O})^{(n-1)})^{K.n})}(\frac{1}{O})^n}{(1 + (\frac{1}{O})^{n-1})^2}$$

In more detail, in the case that $K.n$ **is odd** (and taking K and n as constants), the middle factor in the numerator of $\frac{\partial \mathbb{E}(D)}{\partial O}$ makes the entire fraction to be negative. Additionally, if $K.n$ **is even**, we can still see that the central factor in the numerator is negative as well (with the exception of a single point where it is zero):

$$-1 + (-(\frac{1}{O})^{(n-1)})^{K.n} + K.n(1 + (\frac{1}{O})^{(n-1)})(-(\frac{1}{O})^{(n-1)})^{K.n}$$

$$= -1 + ((\frac{1}{O})^{(n-1)})^{K.n} + K.n(1 + (\frac{1}{O})^{(n-1)})((\frac{1}{O})^{(n-1)})^{K.n} \text{ , because } K.n \text{ is even.}$$

$$= -1 + ((\frac{1}{O})^{(n-1)})^{K.n}(1 + K.n(1 + (\frac{1}{O})^{(n-1)})) \text{ , after factoring.}$$

$$\leq -1 + ((\frac{1}{O})^{(n-1)})^{K.n}(1 + \frac{3}{2}.K.n), \text{ because } (\frac{1}{O})^{(n-1)} \leq \frac{1}{2}.$$

$$\leq 0, \text{ because } 1 + \frac{3}{2}.K.n \leq ((O)^{(n-1)})^{K.n}.$$

Note, for $O = 2$, $K = 1$ and $n = 2$ the equality holds. Naturally, for any other combination where $O > 2$ or $K > 1$ or $n > 2$ the inequality is strict.

Finally, let us motivate the next section by commenting on some expected limitations of Theorem 1: because of the nature of the expectation of a random variable, the information provided by Theorem 1 is not guarded against a large variance of the random variable D. Also, because of the definition of D, its values are not normalized across different opinion trees. This means that (in principle) when sampling opinion trees, we could face a hypothetical situation like the one represented in Figure 2.3. That is, if we take the average (or calculate the expectation) of D, it might result in a positive value, even if most sampled points are negative. In cases like this, it was incorrect to directly imply that the opinion of the first speaker had advantage over other opinions. Naturally, in absence of any other extra information (additionally to the expectation of D), it would be rational to prefer thinking that the first speaker had advantage. **In the next section we show that** when the set O of opinions is large, the (uniform) selection of a single opinion tree leads (with high probability) to a D with a positive value. Clearly, this statement serves as a complement of Theorem 1.

[O,K,n]	Estimates of $\mathbb{E}(D)$	$\mathbb{E}(D)$
[2,1,2]	0.2477	0.25
[3,1,2]	0.2220	0.2222
[4,1,2]	0.1904	0.1875
[5,1,2]	0.1612	0.16
[6,1,2]	0.1397	0.1388
[7,1,2]	0.1225	0.1224
[8,1,2]	0.1088	0.1093
[9,1,2]	0.0985	0.0987
[10,1,2]	0.0899	0.09
[11,1,2]	0.0826	0.0826
[2,2,2]	0.3111	0.3125
[3,2,2]	0.2414	0.2469
[4,2,2]	0.1990	0.1992
[5,2,2]	0.1646	0.1664
[6,2,2]	0.1438	0.1427
[7,2,2]	0.1242	0.1249
[8,2,2]	0.1110	0.1110
[9,2,2]	0.0996	0.0999
[10,2,2]	0.0909	0.0909
[11,2,2]	0.0835	0.0833
[2,3,2]	0.3250	0.3281
[3,3,2]	0.2489	0.2496
[4,3,2]	0.1994	0.1999
[5,3,2]	0.1655	0.1666
[6,3,2]	0.1411	0.1428
[7,3,2]	0.1255	0.1249
[8,3,2]	0.1116	0.1111
[9,3,2]	0.0999	0.0999
[2,1,3]	0.2029	0.2031
[3,1,3]	0.0992	0.1001
[4,1,3]	0.0584	0.0588
[5,1,3]	0.0387	0.0384
[6,1,3]	0.0271	0.0270
[7,1,3]	0.0198	0.0200
[8,1,3]	0.0153	0.0153
[2,2,3]	0.1989	0.1999
[3,2,3]	0.1004	0.0999
[2,3,3]	0.1996	0.2000
[2,4,3]	0.1978	0.1999

Table 2.2: Theoretical $\mathbb{E}(D)$ and its estimates per profile.

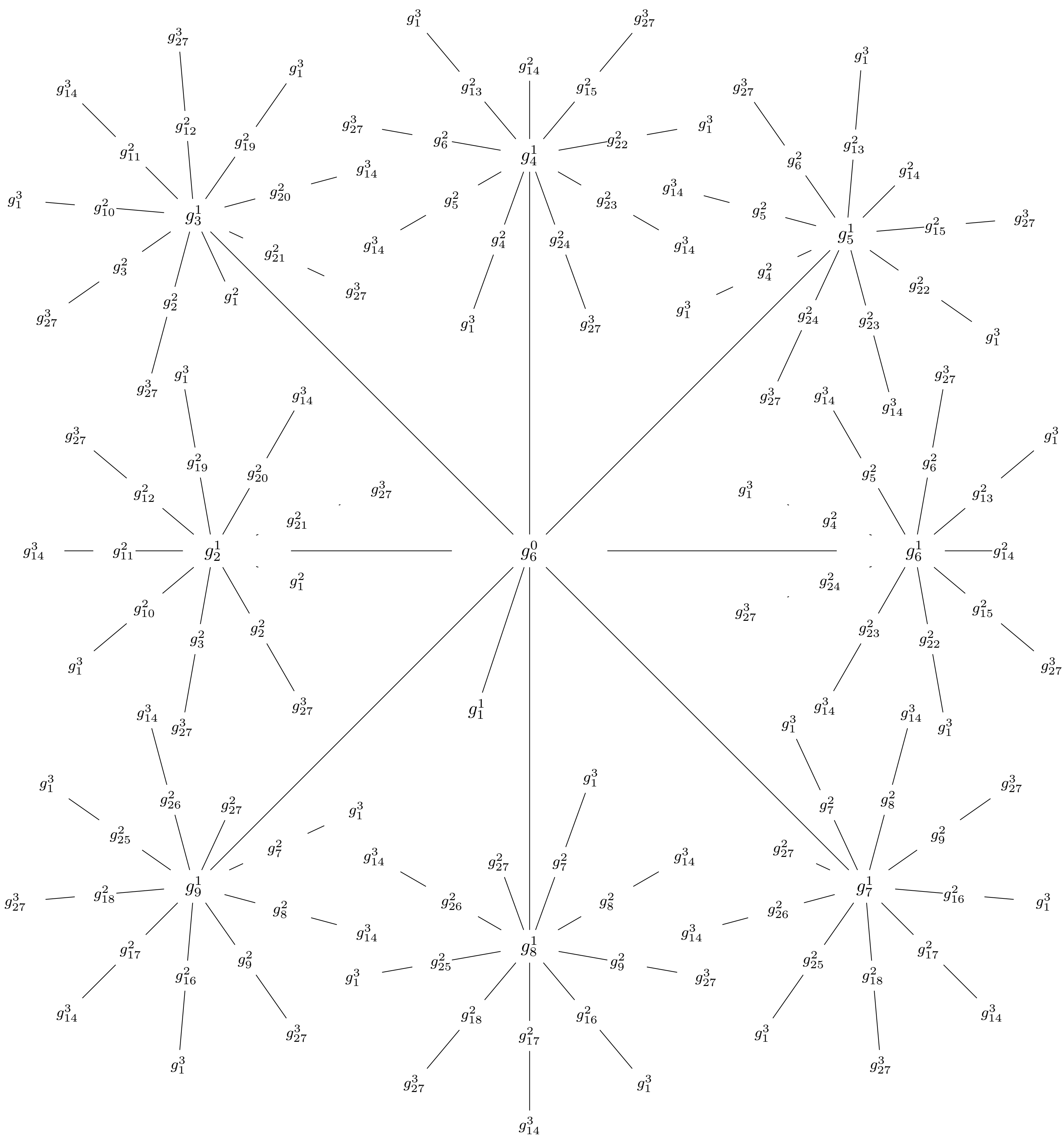

Figure 2.2: Opinion tree with three speakers, one round and three opinions

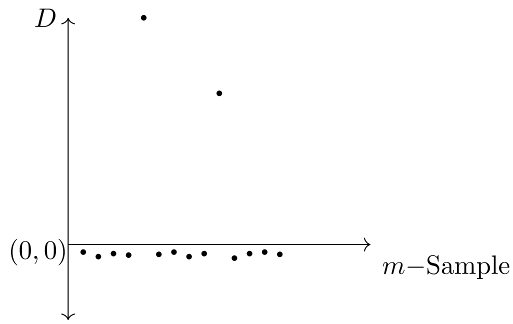

Figure 2.3: Hypothetical sampling of D.

2.3 Probability of anchoring

In this section we consider the scenario in which: (O, K, n) are the parameters of an *opinion tree structure* **and** the cardinality of the set O is large (i.e., we will need to take limits when $O \to \infty$). Additionally, the edge-probabilities for this structure will be selected in a single pick as described in the previous section (i.e., uniformity and independence assumptions are in place). Then, we would like to understand what occurs with the probability of the event $D > 0$ under the given circumstances. In other words, we would like to estimate $P(D > 0)$ in a setting where a single opinion tree is selected uniformly and O is large. Naturally, an answer to this problem (we call it the *Probability of Anchoring Problem*) allowed us to compare the probability of success of the first speaker with the probability of success of any other individual speaker in the selected opinion tree (because of the way in which D is defined and Definition 7).

The intention in this section is to show that under the previous conditions, the value $P(D > 0)$ is very close to one. Given that $D = S(\overrightarrow{0}) - S(\overrightarrow{a})$, $S(\overrightarrow{0}) \geq 0$ and $S(\overrightarrow{a}) \geq 0$, **our basic strategy will be to show that**: there exists a boundary value B such that $P(S(\overrightarrow{a}) \geq B)$ is very small **and** $P(S(\overrightarrow{0}) \leq B)$ is very small as well (see Figure 2.4). Naturally, the existence of B provided us with a "separation point" between $S(\overrightarrow{a})$ and $S(\overrightarrow{0})$ that assured our previous claim about $P(D > 0)$.

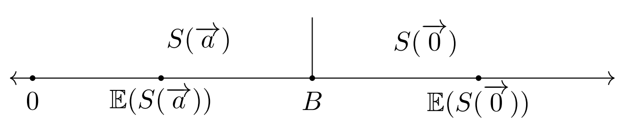

Figure 2.4: A diagram of the boundary of probabilities.

Consequently, we have two statements to show, and we proceed slightly different for each of them. In the first case, for showing that $P(S(\overrightarrow{a}) > B)$ is very small we prove the following inequality:

$$P(|S(\overrightarrow{a}) - \mathbb{E}(S(\overrightarrow{a}))| \geq \frac{\mathbb{E}(D)}{N}) \leq \frac{N \cdot \sqrt{2 \cdot (3 + 2\sqrt{2}) \cdot (2^{\frac{l+1}{2}} - 2)^2}}{O} \tag{2.4}$$

This inequality is just an upper bound on the probability of finding $S(\overrightarrow{a})$ at a distance farther than $\frac{\mathbb{E}(D)}{N}$ from $\mathbb{E}(S(\overrightarrow{a}))$. Bear in mind that because of the definition of D and the linearity of the expectation $\mathbb{E}(D) = \mathbb{E}(S(\overrightarrow{0})) - \mathbb{E}(S(\overrightarrow{a}))$. The parameter N is just a positive integer that indicates the portion of $\mathbb{E}(D)$ that will be considered in the inequality (note, we are using $\mathbb{E}(S(\overrightarrow{a})) + \frac{\mathbb{E}(D)}{N}$ as our B value). As for the bounding term in the right-hand side of the inequality, the most significant fact is that O appears in the denominator. This means that when O is considered arbitrary large and both N and $l = K.n$ are constants, the bounding term is arbitrary small (but positive).

In the second case, for showing that $P(S(\overrightarrow{0}) \leq B)$ is very small, we use the fact that the marginal distributions of a Dirichlet vector $X = X_i$ are Beta distributions. That is, each component X_i of X follows a Beta distribution (see the Dirichlet distribution section of Basic Notions). Because of this, we then know the *cumulative distribution function* (CDF) of the **simplest term** in $S(\overrightarrow{0})$, which **is the one** corresponding to **the edge** that goes **from** the root of the *opinion tree structure* **to** the vertex with label $\overrightarrow{0}$ in the first level of the tree. For instance, the simplest term in Figure 1.1 is p_1; and in Figure 2.2, it is the one corresponding to the edge that goes from g_6^0 to g_1^1. But, **how do we use this fact for our cause** (i.e., that we know the CDF of this term)? Well, we will show that when O is arbitrary large, $x = \mathbb{E}(S(\overrightarrow{a})) + \frac{\mathbb{E}(D)}{N}$, $\alpha_i = 1$ and $\bar{\alpha_i} = O^{(n-1)} - 1$:

$$CDF_{X_i}(x, \alpha_i, \bar{\alpha_i}) = 1 - (\frac{1}{e})^{\frac{1}{N}} \tag{2.5}$$

Naturally, this cumulative distribution function gives us the probability that the random variable X_i takes on a value that is lower than or equal to $\mathbb{E}(S(\overrightarrow{a})) + \frac{\mathbb{E}(D)}{N}$ (bear in mind that we are using this value as our B point). Also, the values α_i and $\bar{\alpha_i}$ are the usual parameters corresponding to a uniform Dirichlet distribution. More important, as we mentioned before N is a constant, but when it takes a fixed large value, the probability value $P(X_i \leq \mathbb{E}(S(\overrightarrow{a})) + \frac{\mathbb{E}(D)}{N})$ is very close to zero (as Eq. 2.5 describes). Further, because of the nature of exponentiation, the same is true without N being particularly large. **Consequently**, if we secure the previous probability bound for the simplest term of $S(\overrightarrow{0})$, the same must hold for $S(\overrightarrow{0})$ itself (recall, there are no negative terms in $S(\overrightarrow{0})$).

Next, we reformulate the former paragraphs in a theorem, it essentially says **that** Eq. 2.4 and Eq. 2.5 hold **and that** the right-hand side of Eq. 2.4 can be made arbitrary small:

Theorem 5. *Provided that the premises of Eq. 2.4 and Eq. 2.5 hold: for all* $\epsilon > 0$ *there exists a real number* $B = \mathbb{E}(S(\overrightarrow{a})) + \frac{\mathbb{E}(D)}{N}$ *such that* $P(S(\overrightarrow{0}) \leq B) \leq 1 - (\frac{1}{e})^{\frac{1}{N}}$ *and* $P(S(\overrightarrow{a}) \geq B) \leq \epsilon$*.*

Before we proceed with the proof of the theorem (which heavily relies on Eq. 2.4 and Eq. 2.5), let us first observe how to **use it**. That is, we are given (O, K, n), a positive ϵ, and we know that O can be arbitrary large (in particular, larger compared with any other parameter involved in this result). Next, we need to produce a separation point B that makes $P(S(\overrightarrow{a}) \geq B)$ **and** $P(S(\overrightarrow{0}) \leq B)$ very small. Accordingly, we select a convenient large enough value for N, and define B as in the theorem. This assures that $P(S(\overrightarrow{0}) \leq B)$ is very small (because N is large). Moreover, given that we can make ϵ be arbitrary small, so will $P(S(\overrightarrow{a}) \geq B)$ be.

As it was mentioned before, the proof of this theorem relies on Eq. 2.4 and Eq. 2.5, and these have been **encoded as Lemma 10 and Lemma 11 respectively**. So, we will have a tidy short proof of the theorem now (see Appendix B for a proof of Theorem 5) and two longer ones for the lemmas later.

As it was the case in the previous section, before we proceed with the proof of the key results (i.e., Lemma 10 and Lemma 11), we need four utility lemmas. They are mostly properties about the first and second moments of $S(\overrightarrow{a})$. Accordingly, we start with a lemma that states an asymptotic expression for $\mathbb{E}(S(\overrightarrow{a}))$.

Lemma 6. *Given an opinion tree structure with parameters* (O, K, n)*, and assuming that the edge-probabilities for* (O, K, n) *are selected under the usual uniformity and independence constraints:*

$$\lim_{O \to \infty} \frac{\mathbb{E}(S(\overrightarrow{a}))}{\frac{\ell - 1}{O^n}} = 1$$

See Appendix B for a proof of Lemma 6.

The second utility result concerns the variance $\mathbb{V}(S(\overrightarrow{a}))$ of $S(\overrightarrow{a})$. It asserts the existence of an upper bound $F(O)$ of $\mathbb{V}(S(\overrightarrow{a}))$, and states the asymptotic behavior of $F(O)$ when O is large.

Lemma 7. *Given an opinion tree structure with parameters* (O, K, n)*, and assuming that the edge-probabilities for* (O, K, n) *are selected under the usual uniformity and independence constraints, there exist a function* $F(O)$ *such that the following holds:*

$$\mathbb{V}(S(\overrightarrow{a})) \leq F(O)$$

$$\lim_{O \to \infty} \frac{F(O)}{\frac{(3+2\sqrt{2}) \cdot (2^{\frac{l+1}{2}} - 2)^2}{O^{2n}}} = 1$$

See Appendix B for a proof of Lemma 7.

In the proof of the previous lemma we use the following upper-bound result:

Lemma 8. *Given an opinion tree structure with parameters (O, K, n), and assuming that the edge-probabilities for (O, K, n) are selected under the usual uniformity and independence constraints, and that E_i and E_k are defined as in the proof of Lemma 7, the following holds:*

$$\mathbb{E}(E_i \cdot E_k) \leq L_i^{A,a} \cdot L_k^{A,a} \cdot \left(\frac{2}{(O^{n-1}+1) \cdot (O^{n-1})}\right)^{\frac{i+k}{2}}$$

See Appendix B for a proof of Lemma 8.

Our last utility lemma is derived from the mixed moments of a random vector that follows a uniform Dirichlet distribution(see the Dirichlet distribution section of Basic Notions).

Lemma 9. *Given an opinion tree structure with parameters (O, K, n) **and** assuming that the edge-probabilities for (O, K, n) are selected under the usual uniformity and independence constraints **and** that p_i and p_j are edge-probabilities that share the same parent vertex (but, $i \neq j$), **it holds that** $\mathbb{E}(p_i^2) > (\mathbb{E}(p_i))^2$ and $\mathbb{E}(p_i^2) > \mathbb{E}(p_i \cdot p_j)$. In more detail:*

$$\mathbb{E}(p_i) = \frac{1}{O^{n-1}}$$
$$\mathbb{E}(p_i^2) = \frac{2}{(O^{n-1}+1)(O^{n-1})}$$
$$\mathbb{E}(p_i \cdot p_j) = \frac{1}{(O^{n-1}+1)(O^{n-1})}$$

See Appendix B for a proof of Lemma 9.

Now, we are ready to proceed with the two important lemmas.

Lemma 10. *Given an opinion tree structure with parameters (O, K, n), and **assuming that**: O is arbitrary large **and** the edge-probabilities for (O, K, n) are selected under the usual uniformity and independence constraints, the following holds:*

$$P(|S(\overrightarrow{a}) - \mathbb{E}(S(\overrightarrow{a}))| \geq \frac{\mathbb{E}(D)}{N}) \leq \frac{N \cdot \sqrt{2 \cdot ((3+2\sqrt{2}) \cdot (2^{\frac{l+1}{2}} - 2)^2)}}{O}$$

See Appendix B for a proof of Lemma 10.

Broadly said, the next (and last) lemma of this section evaluates the cumulative distribution function $CDF_{X_i}(x, \alpha_i, \bar{\alpha_i})$ of a random variable X_i on particular values of its parameters $(x, \alpha_i, \bar{\alpha_i})$. Moreover, it also presents the expression of this evaluation for the particular case in which O is large. Naturally, as we mentioned before, the random variable X_i is a component of a uniform Dirichlet random vector, which entails that X_i follows a Beta distribution. Because of this, we know its cumulative distribution function.

Lemma 11. *Given an opinion tree structure with parameters (O, K, n), and assuming that the edge-probabilities for (O, K, n) are selected under the usual uniformity and independence constraints, it holds that when O is arbitrary large, $x = \mathbb{E}(S(\overrightarrow{a})) + \frac{\mathbb{E}(D)}{N}$, $\alpha_i = 1$ and $\bar{\alpha_i} = O^{(n-1)} - 1$:*

$$CDF_{X_i}(x, \alpha_i, \bar{\alpha_i}) = 1 - (\frac{1}{e})^{\frac{1}{N}}$$

See Appendix B for a proof of Lemma 11.

As we did in the previous section, here we present some simulations too. This time, our motivation comes from the requirement of an arbitrary large O in Lemma 10 and Lemma 11. As we saw, the use of this assumption had two different intentions: **first**, it allowed us to work (in the proofs of the lemmas) with asymptotic expressions that were obtained from limit results (i.e., Lemma 6). **Second**, it guaranteed that $N \cdot \sqrt{2 \cdot ((3 + 2\sqrt{2}) \cdot (2^{\frac{l+1}{2}} - 2)^2)} << O$ holds, and because of this, the bound provided by Eq. 2.4 became more interesting. More important, from Eq. 2.4 we know that the bound (on the right-hand side) is strictly decreasing when O increases. Naturally, in this scenario the previous ideas trigger **the following question**: is it the case that for small sizes of O the probability $P(D > 0)$ increases when O increases? In other words, could we get a similar behavior to the one presented in Theorem 5 even if we did not ask that O was arbitrary large?

For answering the previous question (for some instances of (O, K, n) only), we present simulations of models with profiles that involve small values of O, and we then observe whether in these cases the estimation of $P(D > 0)$ increases when O does it. Consequently, we can see **three columns** in Table 2.3: the profiles of models, the estimates of $P(D > 0)$ and the evaluation of $\sqrt{2 \cdot (3 + 2\sqrt{2}) \cdot (2^{\frac{l+1}{2}} - 2)^2}$ for each particular profile. Regarding the experiments (particulars about the generation of the data can be found in Appendix D.1 (Experiment2)), they were conducted in a similar way to the one explained in the previous section (for estimating $\mathbb{E}(D)$). That is, we create the *opinion tree structure* corresponding to the parameters (O, K, n), and generate a uniform Dirichlet vector of probabilities for each internal vertex. Then, we calculate the value of D, and repeat the same experiment several times. The ratio **of** the number of experiments that resulted in $D > 0$ **to** all outcomes **is our estimator of** $P(D > 0)$. Note, this is a safe way to proceed because we are considering the event $D > 0$. Accordingly, the expectation of the indicator function of this event is equal to the probability of this event (expectation and probability of an indicator function are always equal). Then, the strong law of larger number provides us with a good estimator for $P(D > 0)$ (i.e., the average of the results of the experiments). More important, regarding **the answer to our original question**: the statement "the estimation of $P(D > 0)$ increases when O increases" holds indeed.

Discussion/Conclusion. Next, we revisit the central ideas of the chapter. We also discuss some important points that were left out before (with the intention of keeping the main arguments as clear as possible), but they still need to be considered. As we did in the previous chapter, we use a simple dialogue style here too.

Which one was the leading idea of the chapter? In a situation in which a group of speakers is about to enter a debate (each speaker with a given presenting position and with an opinion on the topic of debate), the opinion of the first speaker has an overall better chance of prevailing than any other opinion of the group. Moreover, the same is true **with high probability** if: in a scenario with many opinions, a single opinion tree was selected (uniformly).

Which arguments were presented to justify the previous idea? **First**, in Definition 7 we saw that a particular opinion tree can be translated into a probability space (so that questions about the outcomes of a debate can be made precise in the safe ground of a probability space). **Second**, Theorem 1 showed (because $\mathbb{E}(D)$ is positive) that for any *opinion tree structure*, if we sampled its edge-probabilities in a uniform and independent form, the opinion of the first speaker has an overall better chance of prevailing than any other opinion of the group. **Third**, in Theorem 5 we presented that: **when** O is arbitrary large, the larger the set O of opinions is, the surer we are **that** when we pick a single opinion tree uniformly, the opinion of the first speaker is the outcome with better prevailing chances. **Fourth**, we demonstrated that not only in the previous case, but also in cases with small sizes of O, the idea "increasing the size of O favours the first speaker" seems to be true. Evidence of this was shown in terms of simulations.

Are the uniformity and independence requirements true to life? Independence and uniformity are assumptions that we kept using through the entire chapter, and from them we profited in key moments (e.g., Lemma 2, Lemma 10, and Lemma 11). Broadly speaking, with these premises we intend to model that: first, when we sample opinion trees (by sampling their edge-probabilities), the selection of probability values **for** edges under different vertices does not affect one each other. Second, when we sample opinion trees, all of them are equally probable to occur. Consequently, when together, these constraints provide us with a "fair" overview of the family of opinion trees that we are dealing with. That is, we wanted to survey the space of opinion trees. **But, is that what happens in our everyday life when a group of individuals is about to engage in a debate** (i.e., does the *opinion tree structure* that describes the debate that is about to start follow the previous two requirements)? **From our perspective, the answer depends on the information that we have or assume about the group in question (i.e., the principle of indifference is not an option here)**. In general, if we consider a known group, we are more sympathetic to a view **in which** some sub regions of a family of opinion trees occur more probably than others (in this case, our current uniformity assumption failed). Which regions exactly, **and** how their probability assignments might look like? This might be contingent on various circumstances, some more stable than others. For instance: the rationality of the speakers, their beliefs, psychological strength, or even their moods might influence a debate. Naturally, in this case (where background knowledge about a group existed), one could attempt to incorporate this knowledge in the distribution function that is used for the sampling of opinion trees. Clearly, for presenting equivalent results to those shown in this chapter, this potentially new setup might require more sophisticated calculations than what was needed in here. **However, the results presented in this chapter are (primarily) about families of opinion trees, and not (directly) about a particular real-life situation that we needed to model**. That is, the ideas presented here are essentially intended to highlight the strong presence of the opinion of the first speaker in the spectrum of opinion trees that one might consider.

Regarding debates, is the idea "the first speaker should have a less strong presence" feasible? If the intention is to reduce $\mathbb{E}(D)$ (because of Theorem 1), we can see that increasing the number of acceptable opinions might be a plausible option. **However, this comes at a price**: because of Theorem 5, we know that allowing a larger O increases the chance that we get an opinion tree where the opinion of the first speaker was the strongest as well (if the opinion tree was selected uniformly). Another option for reducing $\mathbb{E}(D)$ might be (this one with less impact): if O and n are fixed, we could select K in a way such that it is as small as possible and makes $\ell = K \cdot n$ to be an even number.

What is needed for closing the gap between the simulation and the analytical results presented in this chapter? **On the one hand**, the analytical results (based on Lemma 10 and Lemma 11) **assure that**: assuming an arbitrary large O and under the usual uniformity and independence constraints in the pick of an opinion tree, the opinion of the first speaker has a better chance (in that opinion tree) of prevailing than any other opinion of the group. Moreover, this probability increases when O did it. **On the other hand**, the simulation results showed the same pattern of probability increasing (in favor of the first speaker) in particular profiles of (O, K, n) with small values of O. Consequently, an ideal way to close this gap would be: first, one should obtain the cumulative distribution function of D. Second, one should verify that it is a monotonically non-increasing function in O (with K, n fixed).

[O,K,n]	Estimates of $P(D > 0)$	$\sqrt{2 \cdot (3 + 2\sqrt{2}) \cdot (2^{\frac{l+1}{2}} - 2)^2}$
[2,1,2]	0.6958	2.8284
[3,1,2]	0.7678	2.8284
[4,1,2]	0.8205	2.8284
[5,1,2]	0.8483	2.8284
[6,1,2]	0.8709	2.8284
[7,1,2]	0.8874	2.8284
[8,1,2]	0.8952	2.8284
[9,1,2]	0.9087	2.8284
[10,1,2]	0.9120	2.8284
[11,1,2]	0.9232	2.8284
[2,2,2]	0.738	12.4853
[3,2,2]	0.7799	12.4853
[4,2,2]	0.8184	12.4853
[5,2,2]	0.84	12.4853
[6,2,2]	0.8623	12.4853
[7,2,2]	0.8753	12.4853
[8,2,2]	0.8873	12.4853
[9,2,2]	0.8982	12.4853
[10,2,2]	0.9068	12.4853
[11,2,2]	0.9132	12.4853
[2,3,2]	0.7414	31.799
[3,3,2]	0.7821	31.799
[4,3,2]	0.8170	31.799
[5,3,2]	0.8403	31.799
[6,3,2]	0.8612	31.799
[7,3,2]	0.8766	31.799
[8,3,2]	0.8842	31.799
[9,3,2]	0.8952	31.799
[2,1,3]	0.7821	6.8284
[3,1,3]	0.8503	6.8284
[4,1,3]	0.89	6.8284
[5,1,3]	0.9149	6.8284
[6,1,3]	0.9351	6.8284
[7,1,3]	0.9467	6.8284
[8,1,3]	0.9553	6.8284
[2,2,3]	0.7728	31.799
[3,2,3]	0.8421	31.799
[2,3,3]	0.7728	102.426
[2,4,3]	0.7744	302.191

Table 2.3: Estimates of $P(D > 0)$. The computation involved 20000 experiments for each profile.

Chapter 3

Models: particular cases

The ideas presented in the previous chapter have a central role in this dissertation, they concern the impact of the order of speakers on *opinion tree structures*. In this chapter we address a slightly more practical problem, and the intention is that its solution complements what we just learnt about *opinion tree structures*. **The problem itself can be described as follows**: a single model of deliberation is given (as a *black box*), and we need to answer the question—*To which extent does this instance present order dependence?* Naturally, addressing this question would help us to better understand the significance of the order of speakers in individual models of deliberation. **A motivation** for studying models in this individual way can be found in the results of the preceding chapter. That is, **because we know that** when a family of opinion trees is uniformly inspected, the opinion of the first speaker is overall stronger than others, **we would probably like/need to know** whether in a carefully designed model, the opinion of the first speaker was the strongest as well.

Given that models can be produced with the intention of capturing diverse contingent events, **we should not expect that** a general result on particular properties of each individual model **could have** an analytical nature. Accordingly, the ideas that we present here **for studying the order dependence in individual models of deliberation** are based on simulations.

Broadly speaking, there will be **two important moments in this chapter**: first, assuming that a model is not given as a *black box* but in its explicit form instead (i.e., **we have a scenario with perfect information**, like the one expressed in Definition 4), we define precisely the problems that we would like to solve here. Then, in a second moment, **the previous perfect information assumption is turned off** (i.e., the *black box* assumption is in place), and we describe the same problems from a parameter estimation perspective. The reason behind this approach is that we would like to keep our main questions (on order dependence) separated from others that arise from technical difficulties related to the estimation of parameters in the black box case. Naturally, a more detailed motivation for the parameter estimation approach will be provided in the corresponding section.

3.1 The significance of the order of speakers

In an ideal scenario, a model of deliberation is given by a tuple (O, K, n, U, C) as it was described in Definition 4. For instance, in the particular case in which $(O, K, n) = (3, 1, 3)$, we could picture the labels in Figure 3.1 as the roots of **all** opinion trees in U. Note, here the arrangement of labels follows the same order that was suggested in the comments on Figure 1.2. That is, root labels that are permutations of each other share the same row in this description.

$$
\begin{array}{c}
[0,1,0] \quad [1,0,0] \quad [0,0,1] \\
[0,2,0] \quad [2,0,0] \quad [0,0,2] \\
[1,1,0] \quad [0,1,1] \quad [1,0,1] \\
[2,2,0] \quad [0,2,2] \quad [2,0,2] \\
[1,2,1] \quad [2,1,1] \quad [1,1,2] \\
[2,2,1] \quad [1,2,2] \quad [2,1,2] \\
[1,2,0] \quad [2,1,0] \quad [0,2,1] \quad [2,0,1] \quad [0,1,2] \quad [1,0,2]
\end{array}
$$

Figure 3.1: Roots of opinion trees in a model with parameters $(O, K, n) = (3, 1, 3)$.

Also, under the assumption that we have access to the edge-probabilities of each opinion tree in U, we could (for each opinion tree) obtain the probability of success of each consensus outcome. The exact procedure for this calculation was described in the context of Definition 7.

In our particular example, the consensus outcomes are $[0, 0, 0], [1, 1, 1]$, and $[2, 2, 2]$. For them, we specified (in Figure 3.2 and Figure 3.3) the toy probabilities for two different scenarios (denoted A, B and described below). These toy probabilities are just the probabilities of each consensus case **in** opinion trees with root labels in Figure 3.1. In both cases (A and B) each tuple of probabilities has the form $[p_{[0,0,0]}, p_{[1,1,1]}, p_{[2,2,2]}]$. For instance, in the opinion tree with root $[0, 1, 0]$, the probability of the consensus state $[0, 0, 0]$ is 0.6 (in the scenario A, which corresponds to Figure 3.2). Consequently, for the same opinion tree and scenario, the probabilities of the consensus states $[1, 1, 1]$ and $[2, 2, 2]$ are 0.3 and 0.1 respectively.

Now, **an interesting question is**: *what could we say about the significance of the order of the speakers in A and B?* **A first** noticeable aspect about A is that for each row, the probability vectors are exactly the same. Naturally, the same property does not hold for B. Note, the previous circumstance in A is an intentional ideal situation: it means that according to this model, if the speakers were about to start a debate (each speaker has an opinion on a topic), the order in which they presented does not change the odds of the outcome of the debate. Next, we express the difference between the previous scenarios in a more general problem-like form.

Definition 9 (Order Dependence Problem)**.** Given a model of deliberation $M = (O, K, n, U, C)$, the Order Dependence Problem **asks whether** any two opinion trees in U **with** root labels that are permutation of each other **have** the same probability with respect to each consensus opinion x ($x \in O^n$ and $C(x) = 1$). If the answer to the problem is "no", **we say that** the model presents order dependence.

$$\begin{array}{ccc}
[0.6, 0.3, 0.1] & [0.6, 0.3, 0.1] & [0.6, 0.3, 0.1] \\
[0.6, 0.1, 0.3] & [0.6, 0.1, 0.3] & [0.6, 0.1, 0.3] \\
[0.3, 0.6, 0.1] & [0.3, 0.6, 0.1] & [0.3, 0.6, 0.1] \\
[0.3, 0.1, 0.6] & [0.3, 0.1, 0.6] & [0.3, 0.1, 0.6] \\
[0.1, 0.6, 0.3] & [0.1, 0.6, 0.3] & [0.1, 0.6, 0.3] \\
[0.1, 0.3, 0.6] & [0.1, 0.3, 0.6] & [0.1, 0.3, 0.6]
\end{array}$$
$$[\frac{1}{3}, \frac{1}{3}, \frac{1}{3}] \quad [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}] \quad [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}] \quad [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}] \quad [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}] \quad [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$$

Figure 3.2: Scenario A shows the probability of consensus for each plausible outcome. This is a model with parameters $(O, K, n) = (3, 1, 3)$. For each tuple, the format is $[p_{[0,0,0]}, p_{[1,1,1]}, p_{[2,2,2]}]$, and subscripts indicate the consensus opinions.

$$\begin{array}{ccc}
[0.6, 0.3, 0.1] & [0.6, 0.3, 0.1] & [0.6, 0.3, 0.1] \\
[0.6, 0.1, 0.3] & [0.6, 0.1, 0.3] & [0.6, 0.1, 0.3] \\
[0.3, 0.6, 0.1] & [0.3, 0.6, 0.1] & [0.3, 0.6, 0.1] \\
[0.3, 0.1, 0.6] & [0.3, 0.1, 0.6] & [0.3, 0.1, 0.6] \\
[0.1, 0.6, 0.3] & [0.1, 0.6, 0.3] & [0.1, 0.6, 0.3] \\
[0.1, 0.3, 0.6] & [0.1, 0.3, 0.6] & [0.1, 0.3, 0.6]
\end{array}$$
$$[\frac{1}{4}, \frac{2}{3} - \frac{1}{4}, \frac{1}{3}] \quad [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}] \quad [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}] \quad [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}] \quad [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}] \quad [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$$

Figure 3.3: Scenario B shows the probability of consensus for each plausible outcome. This is a model with parameters $(O, K, n) = (3, 1, 3)$. For each tuple, the format is $[p_{[0,0,0]}, p_{[1,1,1]}, p_{[2,2,2]}]$, and subscripts indicate the consensus opinions. Also, the last row was modified with respect to the one in the previous figure, the changes reflect a little advantage for the first speaker.

A second noticeable aspect about scenario A **is that**: if any debate is equally likely to occur, the probability of consensus is the same for each of the three positions in which a speaker could present. Below, we can see the expressions for these probabilities. The subscript ci in p_{ci} stands as a short-name for **the event**: *the initial opinion of speaker i was the consensus of the debate*. Again, the same property does not hold for B (see the expressions for $p^{\star}_{ci}$). Naturally, a model corresponding to scenario A captures a situation in which a speaker would not have a pragmatic reason to prefer a particular position over another for presenting her opinion (even if she had access to the probability values).

$$p_{c0} = \frac{0.6+0.3+0.6+0.6+0.3+0.6+0.6+0.3+0.6+0.6+0.3+0.6+0.6+0.3+0.6+0.6+0.3+0.6+\frac{6}{3}}{24} = \frac{11}{24}$$

$$p_{c1} = \frac{0.3+0.6+0.6+0.3+0.6+0.6+0.6+0.6+0.3+0.6+0.6+0.3+0.3+0.6+0.6+0.6+0.6+0.3+\frac{6}{3}}{24} = \frac{11}{24}$$
$$p_{c2} = \frac{0.6+0.6+0.3+0.6+0.6+0.3+0.3+0.6+0.6+0.3+0.6+0.6+0.6+0.6+0.3+0.3+0.6+0.6+\frac{6}{3}}{24} = \frac{11}{24}$$
$$p^{\star}_{c0} = \frac{0.6+0.3+0.6+0.6+0.3+0.6+0.6+0.3+0.6+0.6+0.3+0.6+0.6+0.3+0.6+0.6+0.3+0.6+\frac{25}{12}}{24} = \frac{133}{12 \cdot 24}$$
$$p^{\star}_{c1} = \frac{0.3+0.6+0.6+0.3+0.6+0.6+0.6+0.6+0.3+0.6+0.6+0.3+0.3+0.6+0.6+0.6+0.6+0.3+\frac{6}{3}}{24} = \frac{11}{24}$$
$$p^{\star}_{c2} = \frac{0.6+0.6+0.3+0.6+0.6+0.3+0.3+0.6+0.6+0.3+0.6+0.6+0.6+0.6+0.3+0.3+0.6+0.6+\frac{23}{12}}{24} = \frac{131}{12 \cdot 24}$$

As it might be expected, these probabilities do not meet the usual **unit measure requirement**. That is, $p_{c0} + p_{c1} + p_{c2} > 1$ and $p^{\star}_{c0} + p^{\star}_{c1} + p^{\star}_{c2} > 1$. The reason for this circumstance is that in most cases (see the roots of opinion trees in Figure 3.1), an initial opinion is shared by more than one speaker. **In what follows**, we use the probabilities p_{ci} as another indicator of the significance of the position of speakers in a model. Note, we could see this as a less strict indicator than the one described just before in Definition 9. Next, we express the difference between the previous scenarios (same p_{ci} vs. different p_{ci}) in a problem-like form.

Definition 10 (Anchoring Problem)**.** Given a model of deliberation expressed as $M = (O, K, n, U, C)$, the Anchoring Problem asks for the probabilities of consensus per position (i.e., each p_{ci}). Moreover, if the value p_{c0} is greater than p_{ct} for each $t > 0$, we say that the model M presents anchoring.

Related to the previous problems, there are four important details that we should keep in mind: **first**, both problems can be solved in an effective way. That is, for a given model $M = (O, K, n, U, C)$, there is a constructive procedure that solves each of these problems. **Second**, in both cases one could extend the problems so that more detailed information is asked. For instance, in the case of the *Order Dependence Problem*, we could instead of asking a "Yes"/"No" question, ask for the ratio $\frac{\text{number of rows with order dependence}}{\text{total number of rows}}$. Naturally, this might be more explanatory than the original binary problem. **Third**, for the investigation of the significance of the order of speakers in a model, the use of one problem or the other depends on how strict do we want to be with respect to the model in question (i.e., the *Order Dependence Problem* is a more strict test). **Fourth**, in the case of the *Anchoring Problem*, **the existence of** two positions s and t such that $p_{cs} > p_{ct}$ holds **implies the existence of** an opinion op such that: **the probability of the event** '*the initial opinion of the speaker with position s is op and op is the consensus of the debate*' **is greater than the probability of the event** '*the initial opinion of the speaker with position t is op, and op is the consensus of the debate*'. Clearly, this is an unwelcome property (i.e., that the probability of consensus of an opinion changes with the change of the speaking position), and it serves as a motivation for making an active use of the *Anchoring Problem* as a test for models. As an instance of this situation, we could take from scenario B the case of $p^{\star}_{c0} > p^{\star}_{c2}$. In this example, for $s = 0, t = 2$ and $op = 1$, we obtain the following probabilities for the mentioned events:

$$p^{1}_{c0} = \frac{0.3 + 0.6 + 0.6 + 0.6 + 0.6 + 0.3 + \frac{5}{12} + 0.3}{24}$$

$$p^{1}_{c2} = \frac{0.3 + 0.6 + 0.6 + 0.6 + 0.6 + 0.3 + 0.3 + 0.3}{24}$$

So far, given a model M, **we have assumed** a scenario with perfect information in which (for each opinion tree of M) there is access to the probability of each consensus opinion. In the next section, **we turn that assumption off**, and discuss how could we continue profiting from the solutions of the problems that we just presented in order to investigate the significance of the order of speakers in an individual model of deliberation.

3.2 Estimation of the order dependence

In this brief section we discuss how to approach the *Order Dependence Problem* and the *Anchoring Problem* from a more pragmatic perspective (i.e., one in which we do not have direct access to the probabilities of consensus that resulted from the individual opinion trees of a model). **The motivation to do this** is that when a model is generated, it is usually described using a mixture of natural language and mathematical rules/statements. Consequently, under these circumstances it is often complex to have the required regularity for the derivation of literal probability values corresponding to individual consensus opinions. However, even if this was the case (i.e., we have a *black box* environment), we were still interested in having information about the significance of the order of the speakers in a given model.

In the rest of the section, **we regard each opinion tree of a model as a black-box** that produces consensus opinions (i.e., a model is a *black box* itself). Moreover, each consensus opinion is produced with a certain probability that is unknown to us, but it is assumed to be consistent with Definition 7. **For instance**, if we consider the opinion trees with root labels in Figure 3.1, we could picture their probabilities of consensus opinions as in Figure 3.4(again, their values are unknown to us). Next, we can notice that **these premises allow for simulations**. That is, we are **able to simulate** any opinion tree of a model by (repeatedly) **asking for** a consensus opinion **from** the black box that mimics the mentioned opinion tree.

$$
\begin{array}{c}
[p_{11}, p_{12}, p_{13}] \quad [p_{14}, p_{15}, p_{16}] \quad [p_{17}, p_{18}, p_{19}] \\
[p_{21}, p_{22}, p_{23}] \quad [p_{24}, p_{25}, p_{26}] \quad [p_{27}, p_{28}, p_{29}] \\
[p_{31}, p_{32}, p_{33}] \quad [p_{34}, p_{35}, p_{36}] \quad [p_{37}, p_{38}, p_{39}] \\
[p_{41}, p_{42}, p_{43}] \quad [p_{44}, p_{45}, p_{46}] \quad [p_{47}, p_{48}, p_{49}] \\
[p_{51}, p_{52}, p_{53}] \quad [p_{54}, p_{55}, p_{56}] \quad [p_{57}, p_{58}, p_{59}] \\
[p_{61}, p_{62}, p_{63}] \quad [p_{64}, p_{65}, p_{66}] \quad [p_{67}, p_{68}, p_{69}] \\
[p_{71}, p_{72}, p_{73}][p_{74}, p_{75}, p_{76}][p_{77}, p_{79}, p_{79}][p_{7\_10}, p_{7\_11}, p_{7\_12}][p_{7\_13}, p_{7\_14}, p_{7\_15}][p_{7\_16}, p_{7\_17}, p_{7\_18}]
\end{array}
$$

Figure 3.4: Model with parameters $(O, K, n) = (3, 1, 3)$. For each opinion tree, the probabilities of consensus are unknown. Each tuple has the format $[p_{[0,0,0]}, p_{[1,1,1]}, p_{[2,2,2]}]$, and subscripts indicate the consensus opinions.

Besides our particular example, **the general challenge in this kind of situations will be**: how could we estimate the probabilities of consensus p_{ij}, so that we can estimate the answers of both the *Order Dependence Problem* and the *Anchoring Problem*. Fortunately, this type of problem has been extensively studied under the subject of *statistical analysis of the output data from a simulation model* (see Rubinstein and Kroese 2016, Ch. 3, 4).

A standard approach to our previous challenge would be as follows: given an opinion tree T in the row i of a model, let us consider the drawing of a random sample of consensus opinions $X_1, X_2, \cdots, X_N$ (obtained via simulations of T). Then, an estimation of a p_{ij} **for T** can be obtained as $\hat{p_{ij}} := \frac{1}{N} \cdot \sum_{l=1}^{N} I^{i,j}_{X_l=C_{ij}}$. In this expression, the term $I^{i,j}_{X_l=C_{ij}}$ is an indicator function that tells whether the sample X_l is equal to the consensus opinion C_{ij}, which is the consensus opinion corresponding to the probability value p_{ij}. **For instance**, if we consider the opinion tree with root label $[0, 1, 0]$ (see Figure 3.1 and Figure 3.4). The estimators corresponding to the probabilities p_{11}, p_{12}, p_{13} are as follows:

$$\hat{p_{11}} = \frac{1}{N} \cdot \sum_{l=1}^{N} I^{i,j}_{X_l=[0,0,0]}$$

$$\hat{p_{12}} = \frac{1}{N} \cdot \sum_{l=1}^{N} I^{i,j}_{X_l=[1,1,1]}$$

$$\hat{p_{13}} = \frac{1}{N} \cdot \sum_{l=1}^{N} I^{i,j}_{X_l=[2,2,2]}$$

A justification of the previous procedure can be described in two parts: **first**, the nature of the indicator functions **assures that** the probabilities p_{ij} of the event $\{X_l = C_{ij}\}$ **and** the expectations $\mathbb{E}(I^{i,j}_{X_l=C_{ij}})$ are identical. **Second**, the strong law of large numbers **assures that** $\hat{p_{ij}}$ converges to $\mathbb{E}(I^{i,j}_{X_l=C_{ij}})$ –therefore, to p_{ij} as well– with probability 1 as $N \to \infty$.

Naturally, after obtaining the estimates of the consensus probabilities for each opinion tree of a model, **the initial task is not yet completed**. We still needed to produce estimates for the solutions of the *Order Dependence Problem* and the *Anchoring Problem*. This is a processes that will highly depend on the particular model in question (e.g., on the number of opinions and speakers). But, as we saw in the previous section (where the problems were defined), **it basically involves additions and comparisons of the estimates** $\hat{p_{ij}}$. Therefore, instead of describing particular cases of individual models, here we highlight **two** general aspects that are relevant to most cases: **first**, the previous proposal for estimating the probabilities p_{ij} falls in the category of the "crude" Monte Carlo method (see the Monte Carlo section of Basic Notions for a more extensive example on its use for the calculation of integrals). Weak points of this method are well known, so the approach presented above should be taken as a simple description of a general solution. Consequently, in the analysis of particular models more challenging situations might arise (e.g, the values p_{ij} might be very small), and specific Monte-Carlo-solutions might be needed. **Second**, once we enter the realm of estimators, it is important to keep in mind that certainty is lost, and whenever we needed to state **how close** a p_{ij} is to an estimate $\hat{p_{ij}}$, the appropriate uncertainty and confidence values should be part of the answer too. Moreover, given that for estimating the solution of the *Order Dependence Problem* and the *Anchoring Problem* some operations (mainly additions and comparisons) are going to take place among estimated quantities, it is important to be aware about how the propagation of uncertainties takes place in individual cases (Taylor 1997 **and** Hughes and Hase 2010 provide a pleasant introduction to the topic of error analysis and uncertainties).

Discussion/Conclusion. This chapter was not dedicated to the presentation of analytical or experimental results on deliberative situations. Instead, we have defined and discussed two generic problems, which can be instantiated and solved in the context of individual models of deliberation. More importantly, in the circumstances of each model, the solutions of these problems provide information regarding the impact of the order of speakers on the outcome of deliberations. Naturally, it is crucial to keep always in mind that the nature of this information is probabilistic.

Chapter 4

Alternative models

This chapter can be seen as a reaction chapter. That is, in previous sections we learnt that some unwelcome properties hold true in the world of models of deliberation. Consequently, we react to these circumstances next: we will explore some natural ideas that might help to find better structures of models of deliberation.

In more detail, in the following sections we explore **four alternatives**, and they all share the same basic idea: they are modifications of the initial notion of what we understood as a model of deliberation (see Chapter 1). A detail regarding notation, in the rest of the current chapter, instead of the long term "modification of the initial notion of model of deliberation" we often use the abbreviation *model modification*. **In the first case**, the modification is easy to describe; during a deliberation, the order in which the speakers present their opinions is not going to be fixed anymore. That is, after a speaker spoke, the next one will be always selected at random (uniformly). **A second modification** explores the idea of *agreement by opinion reduction*. In this scenario, a group only decides between two different opinions at a time, and the "winner opinion" will be kept alive (for future rounds) while the "loser" is disregarded at once. **In a third modification**, we carefully restrict the chance of agreements that might arise immediately after the first speaker presented her opinion. Our justification for testing this strategy is purely intuitive: super quick agreements are unlikely in real-life. **The fourth and last** modification explores the option of generating cliques of opinions before the actual deliberation starts. Broadly speaking, this strategy is a form of *opinion reduction* too, and it might happen prior to the beginning of a debate. That is, instead of having complex deliberative situations with a potentially large number of initial opinions, deliberations will take place in a simplified scenario with a smaller number of opinions to defend. Clearly, each of these opinions might be seen as a representative of its clique.

Regarding the analysis of our exploration, in the **first** and the **third** modifications, we directly evaluate (via simulations) the impact of the first speaker's opinion on the outcome of deliberations. Here, we will follow the same strategy as in Chapter 2. That is, we will sample opinion trees and produce an estimate of $P(D > 0)$. Moreover, we will contrast these results with those presented in Table 2.3. Because of the nature of the **second** and the **forth** modifications, their analysis will be slightly different (descriptions will take place in their respective sections).

4.1 Volatile speaking positions

In this section, we modify the basic notion of model introduced in Chapter 1. In particular, we remove the idea that promises a fixed speaking position, which was encoded as follows:

"... The speakers enter the deliberation with an initial opinion each, and what follows respects this dynamics: the first speaker presents her opinion, then all the others update theirs. Next, the second speaker presents his opinion, and all the others update theirs. When the last speaker spoke and all the others updated, the previous presenting/updating process is repeated ... ."

The intention behind this modification is that perhaps (without a fixed order in place) it could provide us with an alternative notion of deliberation in which the opinion of the first speaker did not play the disproportionate role that was recognized in previous sections. Naturally, if we eliminate the idea of a fixed speaking position, we still need to find a way for deciding the next speaker in a debate. Under these circumstances, a random selection (from a uniform distribution) appears to be a plausible contender because it seems to avoid any commitment to the position of the speakers. Broadly said, what we are suggesting is that after a speaker spoke, we should roll a (large enough) dice for deciding the next speaker of a deliberation. Of course, an important detail in this approach is that we must disallow the situation in which a speaker that just presented could speak immediately again. This technicality regarding representation means that if an individual repeated her argument twice, it will be represented only once in the *opinion tree* of a deliberative situation.

As we did in Chapter 1 (with Figure 1.1), here we can use Figure 4.1 to describe the intended behavior of our new *model modification*. Because most of the features of the original notion of model remain intact and only the form of selection of the next speaker changed, the key element in our description is the position of the "*" symbol. As usual, it will occurs in the labels of internal vertices of the opinion tree. The contrast between the two mentioned *opinion trees* is illustrative: **first**, in Figure 1.1, the root vertex $[0^*, 0, 1]$ (which describes the speakers' initial opinions) allows for four transitions after the first speaker spoke. These transitions are $(0, 0^*, 0), (0, 0^*, 1), (0, 1^*, 0)$, and $(0, 1^*, 1)$. In Figure 4.1, starting with the same opinions leads to four transitions as well, and they are $(0, 0, 0^*), (0, 0, 1^*), (0, 1, 0^*)$, and $(0, 1, 1^*)$. Unsurprisingly, here the symbol "*" appears in the third position because in the random choice, the third position was the one selected (in this example). **Second**, in Figure 1.1, we could find the path $[0^*, 0, 1] \to (0, 1^*, 1) \to (0, 1, 1^*) \to (1^*, 1, 1)$ in which the positions of "*" indicate that a linear order of the speakers was respected. However, in Figure 4.1, the labels in the same path are $[0^*, 0, 1] \to (0, 1, 1^*) \to (0, 1^*, 1) \to (1^*, 1, 1)$. Clearly, the difference between these paths reflects that in the second example, the selection of the next speaker was performed randomly as described above. A visible **third difference** between Figure 1.1 and Figure 4.1 is that not every path that appears in the first tree appears in the second one too (the other way around is true as well). However, this situation was expected because it would have been a surprise that a strategy with a "random selection of the next speaker" could easily mimic a sequentially generated opinion tree.

Next, one might expect that in the following lines we could continue with a positive narrative in which we reported that the current *model modification* has a positive impact on deliberations. However, in Table 4.1 we can verify that this is (unfortunately) not the case (particulars about the generation of the data can be found in Appendix D.1 (Experiment3)). That is, when we sample opinion trees like the new modification dictates, we obtain nearly the same results that were obtained in Chapter 2 (Table 2.3). Naturally, this outcome is **counter-intuitive**, and it is a difficult one to anticipate too. So far, we have obtained results that support the idea that the order of the speakers is important for the conclusion of a deliberative situation, and yet, after taking random picks explicitly, the results are roughly the same as they were before.

Fortunately, as it happens in other cases of counter-intuitive scenarios, there is a satisfactory explanation for this one too. This explanation will be provided/discussed later in this chapter in the context of another *model modification*. But, broadly said, what occurred here was that even if we took random picks whenever possible (after a debate started), the chance of agreement immediately after the very first speaker spoke is still in place (and it has a strong impact too). Note that a solution is not as simple as "picking the first speaker randomly as well". If we did so, an uncertainty about "*who is the favored speaker?*" was introduced, but the essential problem persisted: the existence of a speaker with a disproportionate impact on a deliberative process (and only because he was the first to talk . . .).

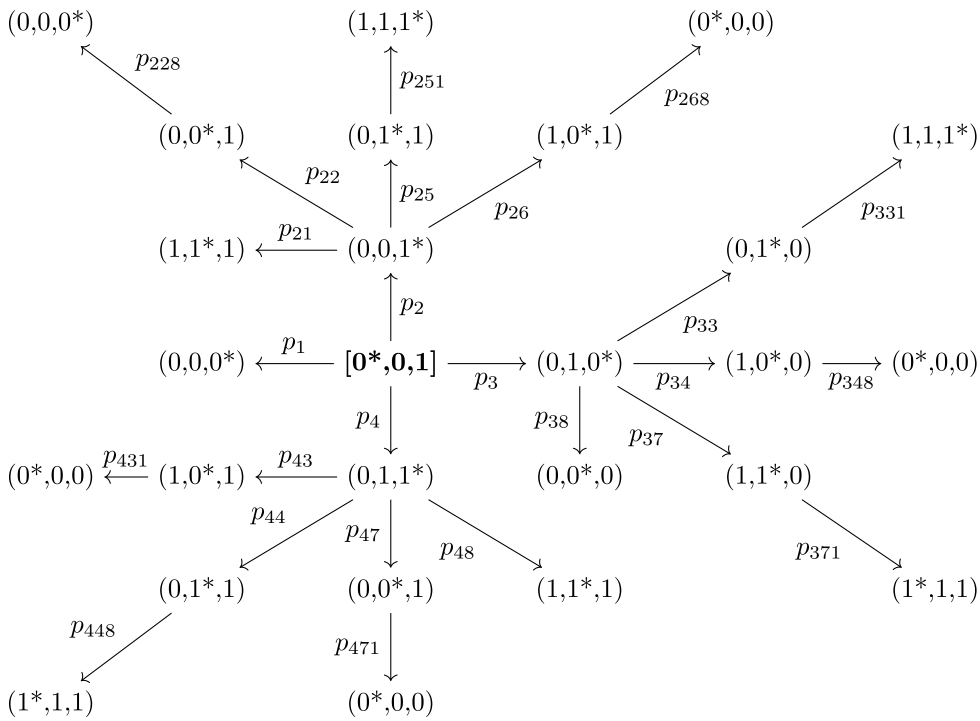

Figure 4.1: Opinion tree with three speakers, one round and two opinions. The selection of the next speaker is random (uniform).

[O,K,n]	Estimates of $P(D > 0)$
[2,1,2]	0.6947
[3,1,2]	0.7748
[4,1,2]	0.8179
[5,1,2]	0.8441
[6,1,2]	0.8724
[7,1,2]	0.8870
[8,1,2]	0.8979
[9,1,2]	0.9071
[10,1,2]	0.9137
[11,1,2]	0.9221
[2,2,2]	0.7359
[3,2,2]	0.7823
[4,2,2]	0.8171
[5,2,2]	0.8430
[6,2,2]	0.8657
[7,2,2]	0.8738
[8,2,2]	0.8881
[9,2,2]	0.8972
[10,2,2]	0.9058
[11,2,2]	0.9134
[2,3,2]	0.7452
[3,3,2]	0.7861
[4,3,2]	0.8149
[5,3,2]	0.8385
[6,3,2]	0.8626
[7,3,2]	0.8749
[8,3,2]	0.8874
[9,3,2]	0.9038
[2,1,3]	0.7767
[3,1,3]	0.8424
[4,1,3]	0.8892
[5,1,3]	0.9172
[6,1,3]	0.9365

Table 4.1: Estimates of $P(D > 0)$. The computation involved 20000 experiments for each profile.

4.2 Agreement by opinion reduction

In this section, we present a second *model modification.* The intention behind this modification is clear: we would like to avoid debates about many opinions at the same time. Instead, we focus in deliberations that must decide (perhaps multiple times) between two opinions only. Naturally, behind this intention, there is still the hope that by modifying the original structure of deliberations, one might be able to reduce the impact of the first speaker's opinion on the outcome of debates.

Next, we illustrate the previous intention with an example: let us assume that Figure 4.2 represents a scenario in which a group of individuals would like to deliberate about the opinions A,B,C and D (different one another). For now, the size of the group is still irrelevant. Interestingly, instead of deliberating under the standard dynamics described in Chapter 1, the group focuses on a debate around the two opinions A and B exclusively. Now, following the description in Figure 4.2, let us assume that the consensus of this partial debate was B. Consequently, the group decides to deliberate about the opinions C and D in the next step, and let us assume that here the consensus was C. Finally, our group reached a situation in which they can deliberate about the opinions B and C, and one might argue that if a consensus was reached (in this case was B), it would be a plausible solution to the original group problem: organize a deliberation about A,B,C and D in order to pick one of them as a consensual opinion.

Naturally, the previous example can be generalized to an arbitrary number of opinions. However, there are still **three details** that needed further clarification before we are ready to show some (simulated) tabular results for different deliberative profiles. **First**, even if the previous structure of deliberations is "new", all particular debates about two opinions will still be conducted under the standard dynamics introduced in Chapter 1 (in our example, they were (A vs. B), (C vs. D), and (B vs. C)). Moreover, in every deliberation about two opinions (e.g., (C vs. D)), we will assume that the opinion of the first speaker is the leftmost one (in this case was C). This requirement allows that nearly any contender opinion might enjoy the potential opportunity of being the opinion of the first speaker at some point during a deliberation. In terms of our example, with the exception of A (which will always be the opinion of the first speaker as long as it prevailed) and D (which will never be the opinion of the first speaker), any other opinion had the chance of being the first speaker's opinion. Of course, a similar idea holds for more than four opinions too.

Second, one might notice that (compared to the original notion of model) in our current *model modification*, it is less simple to **directly** quantify the impact of the opinion of the first speaker. The reason for this circumstance is that several deliberations are in place here (e.g., (A vs. B), (C vs. D), and (B vs. C)) and not just a single one as before. A first potential option for a measure could be to simply aggregate the impact on each "*small deliberation*" that takes place. But, this idea can be quickly ruled out **because** from previous sections (see Table 2.3), we know that the opinion of the first speaker is still too strong in two-opinions debates as well. Consequently, the results of this form of aggregation were both easy to predict and misleading. Accordingly, we take a second alternative here, which is to track the victories of opinions by their initial position. Naturally, this approach includes interesting information about both the strength of the first speaker's opinion and the strength of the last speaker's opinion (via the victory record of the first and the last opinions of the list; A and D in our example). Moreover, statistical information about the victory record of other opinions is intriguing as well because it will show the actual **impact of** taking turns on the role of being the first speaker's opinion **on** the outcome of a debate.

Third, in a scenario in which the initial number of opinions is not a power of two, some extra rules are still needed for making sure that the illustrated process works fine. But, because our primary interest in this chapter is exploratory, we will assume that the number of initial opinions is of the form 2^k for some natural number k, so that we can continue with our investigation.

The next natural step in our exploration is to simulate the previous dynamics on different deliberative profiles **and** examine the collected statistics regarding the strength of the first speaker's opinion (particulars about the generation of the data can be found in Appendix D.1 (Experiment4)). The simulation results are presented in Table 4.2. They are organized in two columns, and as usual, the first one contains the profiles that were inspected. The second column shows the number of "deliberative victories" per initial position of opinions. For instance, if with the profile $[O, K, n] = [4, 1, 2]$, the opinions to discuss were A, B, C, D, the results show that out of 20000 experiments, the debates ended **9691 times** with an A consensus, **4316 times** with a B consensus, **4140 times** with a C consensus, and **1853** times with a D consensus.

Unfortunately, the Table 4.2 does not show encouraging results. That is, we have examined **eight profiles** with **two different** numbers of opinions, **three different** numbers of rounds, and **two different** numbers of individuals. In these profiles, the following ideas seem to hold: first, for each profile, the impact of the first speaker's opinion is still too strong. Second, with the values of **K and n fixed**, the strength of the first speaker's opinion decreases when O increases. Third, with **O and K** fixed, the impact of the first speaker's opinion increases when n increases. Fourth, with **O and n** fixed, the impact of the first speaker's opinion increases when k does it. Moreover, simulations on the extra profile $[O, K, n] = [128, 1, 3]$ suggest that a violent increase of the number of opinion does not look like a positive alternative either:

$$[3573, 982, 989, 279, 1025, 306, 273, 80, 935, 291, 276, 80, 271, 81, 84, 17, 992, 259, 283, 83, 270, 79, 63, 29, 285, 79, 85, 21, 67, 17, 22, 8, 974, 259, 264, 88, 282, 81, 83, 33, 269, 71, 75, 17, 60, 17, 23, 5, 309, 87, 86, 18, 72, 21, 24, 8, 72, 19, 22, 6, 24, 4, 7, 4, 1021, 292, 293, 76, 289, 81, 88, 23, 288, 73, 88, 18, 85, 32, 16, 7, 265, 84, 84, 21, 77, 17, 23, 6, 72, 14, 20, 4, 22, 5, 7, 3, 260, 87, 83, 12, 101, 18, 30, 7, 84, 13, 21, 7, 11, 4, 6, 2, 66, 22, 19, 12, 16, 5, 8, 2, 21, 9, 4, 2, 2, 3, 1, 0]$$

Naturally, it is unfortunate that the current modification did not show positive results either. But, in the next section we present a third modification, and the new report will look better.

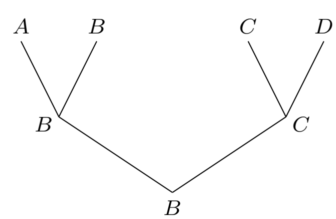

Figure 4.2: Diagram of an opinion tree structure with sequential opinion reduction.

4.3 Limitation on early agreements

In this section, we present a third *model modification*. That is, here we **restrict** the chance of (emergence of) those agreements that might arise immediately after the first speaker presented her opinion for the first time in a debate. Naturally, this modification

[O,K,n]	Debate victories per initial position of opinions
[4,1,2]	[9691, 4316, 4140, 1853]
[8,1,2]	[6697, 2944, 2943, 1360, 2880, 1295, 1292, 589]
[4,2,2]	[10793, 3971, 3870, 1366]
[8,2,2]	[7858, 2970, 2801, 1018, 2862, 1027, 1082, 382]
[4,3,2]	[11141, 3835, 3765, 1259]
[8,3,2]	[8396, 2836, 2787, 959, 2759, 966, 980, 317]
[4,1,3]	[12269, 3452, 3342, 937]
[8,1,3]	[9489, 2703, 2696, 742, 2658, 746, 740, 226]

Table 4.2: For each profile, the table shows a record of debate victories per initial position of opinions. The computation involved 20000 experiments for each profile.

can be justified without effort. One might argue that in real-life deliberative scenarios, humans are not particularly complaisant.

On the model side, *opinion trees* always follow a precise structure, which is easy to describe as a two-element object: **the first one** is an edge that starts at the root of the tree and ends in a consensus vertex. **The second one** is the rest of the opinion tree. In other words, Figure 4.3 can be seen as a template for opinion trees. Consequently, (said again in terms of Figure 4.3), in this section we would like to explore the impact of P_{bound} on the strength of the first speaker's opinion. That is, here we present simulation results for different deliberation-profiles and values of P_{bound} (particulars about the generation of the data can be found in Appendix D.1 (Experiment5)). As usual, next we will draw some conclusions based on this information too.

The outcomes of simulated experiments are presented in Table 4.3. The style of this table is very similar to the one of Table 2.3. That is, because we have sampled *opinion trees* for different deliberative profiles, we can present estimates of $P(D > 0)$ for each of these profiles. Also, we have experimented with multiple values of P_{bound}, and they are included in the table too (per profile). Moreover, these P_{bound} values were intentionally taken "around" the **expected value** of the edge in question, so that **it** provides us with a natural reference point (from Chapter 2 we have an analytic expression for this expectation). Consequently, the expectation values of the edge-probabilities are included in the table too. **For instance**, in the case of profile $[O, K, n] = [2, 1, 2]$ with expected value $\mathbb{E}(p_e) = 0.5$ for the edge-probability, we have experimented with P_{bound} values equal to $0.25, 0.375, 0.625$, and 0.75 respectively.

Naturally, the next interesting question is "How do the $P(D > 0)$ values from Table 4.3 compare to the $P(D > 0)$ values from Table 2.3?". For answering this question, there are two key observations that we would like to highlight. **First**, in Table 4.3, there are profiles in which for some values of P_{bound}, the estimate of $P(D > 0)$ is greater than the corresponding ones (i.e., same profile) from Table 2.3. For instance, for $[O, K, n] = [3, 1, 2]$ with $P_{bound} = 0.4167$, we obtained the estimate of $P(D > 0) = 0.7712$, which is greater than the estimate of $P(D > 0) = 0.7678$ that was obtained for the same profile in Table 2.3. Clearly, there are other profiles with the same behavior. **Second**, (in Table 4.3) if we only concentrate in rows with smaller P_{bound} values than their corresponding expectation values, then in these rows we obtain smaller estimate of $P(D > 0)$ fields than in the same profiles from Table 2.3.

For instance, for $[O, K, n] = [3, 1, 2]$ and $P_{bound} = 0.25$ (which is smaller than $\mathbb{E}(p_e) = 0.3333$), we obtained the estimate of $P(D > 0) = 0.7102$, which is smaller than the estimate of $P(D > 0) = 0.7678$ from Table 2.3. As expected, the same behavior could be observed if P_{bound} is further decreased (e.g., the same profile $[O, K, n] = [3, 1, 2]$ and $P_{bound} = 0.1667$).

The previous analysis suggests that (on the model side) we do have something positive to report. That is, by setting P_{bound} to a smaller value than $\mathbb{E}(p_e)$ (for each particular profile), one could indeed decrease the impact of the first speaker's opinion. Also, let us recall that this idea is relevant to two other interesting moments in this dissertation. **The first one** happened at the conclusion of Chapter 2. Back then, we discussed the style in which *opinion trees* were sampled (i.e., the adequacy of independence and uniformity was challenged). In this regard, a highlighted alternative (to our sampling method) was that **if** we had more information about a specific family of deliberative situations, **then** we could sample in a different way for that particular case. Well, what we presented just before can be seen as form of sampling in which some extra information was assumed (i.e., that deliberations do not end quickly in consensus). **The second moment** occurred in an early section of the current chapter in which a specific *model modification* was explored. The key idea of this modification is that in a deliberative situation, the "next speaker" is always selected at random. In that context, it was unexpected that the opinion of the first speaker was still too strong. But, we just learnt about a factor (i.e., P_{bound}) that was still playing a strong role during a random selection of the next speaker, and this clarifies the previous unexpected circumstances.

Now, even if it seems like the current *model modification* meets our requirements, this is not quite true. Consequently, before ending this section, we would like to signal two details that we should keep in mind regarding this modification. The first is that before applying this notion of model to a real-life scenario, one needs to make sure that it represents/describes the target circumstances properly. In this case, the most important "reality" constraint is that deliberations do not end quickly. The second detail is more technical. That is, one might think that by decreasing P_{bound}, our original problem (i.e., the excessive strength of the first speaker's opinion) was totally tuned. But, this is not the case. In other words, one could set P_{bound} to be equal to a very small value, and it would still be possible that $P(D > 0)$ increases for different profiles. This phenomenon is an echo of Theorem 5, and it can be perceived it in Table 4.3 too. For instance, if we consider $P_{bound} = 0.25$ and the profiles $[2, 1, 2], [3, 1, 2]$, we obtain 0.2885 and 0.7102 as estimates of $P(D > 0)$. Additionally, even for the profile $[4, 1, 2]$ and a more restricted $P_{bound} = 0.125$, we obtain $P(D > 0) = 0.7244$ (which is greater than 0.7102). The key idea here is that a decreasing of P_{bound} is not a universal solution, but it depends on the profile in question.

4.4 Cliques of opinions

In this brief section, we discuss a fourth *model modification*. The hope (here too) is that some changes in the structure of our original notion of model might decrease the impact of the opinion of the first speaker on the outcome of deliberations. This is the last *model modification* that we analyze in this chapter, and there will be no simulations here. Instead, we rely on analytical results that were presented in Chapter 2.

Broadly said, the deliberative situations that we confront in this section are those

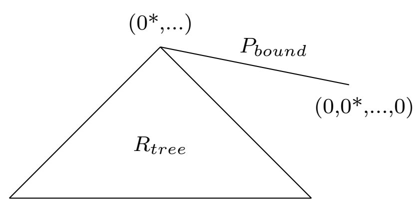

Figure 4.3: The diagram of an opinion tree structure shows a bounded probability of early agreements.

associated to scenarios **in which** the individuals of a group ruled out some of the initial opinions **before** the deliberation even started (here, "deliberation" refers to our standard notion from Chapter 1). However, we are not going to be specific about the reasons that the individuals might have had to proceed in this way (e.g., simplicity, pragmatism, epistemic, etc.). Likewise, we do not name any particular process that a group might have followed to eliminate initial opinions (natural candidates for this role might be *voting systems* or *clustering techniques*). That is, our exploration starts once the opinions have been already removed.

Accordingly, in order to analyze this *model modification* (from our *order dependence* perspective), it is quite natural to select the following as our leading question: *Is it "better" to deliberate about many opinions* ***or*** *about few of them instead?* This question is closely related to another one that was briefly discussed at the *end* of Chapter 2. Next, we will try to reduce the new one to the old one: *Regarding debates, is the idea "the first speaker should have a less strong presence" feasible?* Back then, the argumentation was conducted in terms of O and D (see Definition 8), which are the number of opinions in a debate **and** the difference of strength between the opinion of the first speaker and the opinion of any other speaker respectively. Also, let us recall that our reaction to that question included two ideas: **first**, if O increases then $\mathbb{E}(D)$ decreases. **Second**, in deliberative situations with a large O, the value $P(D > 0)$ is close to the unit.

Clearly, with the previous information at hand, it becomes uncomplicated to answer the first question. That is, in an scenario in which a **single** deliberation is about to start, it would be ideal to avoid that $P(D > 0)$ is near to the unit. Consequently, whenever possible, it would be recommended to decrease the number of opinions before a deliberation starts. Moreover, Table 2.3 suggests that this advice might be valid not only for large values of O but for smaller instances too.

This is an interesting moment because it seems that for the current *model modification* there will be good news only, but there is more to it. **First**, even if decreasing the number of opinions seems to be a plausible option, we should remember that Table 2.3 shows that for profiles with small instances of O, the opinion of the first speaker is too strong anyway. **Second**, even if reducing the number of opinions is a reasonable approach, we should keep in mind that this process/problem might be as complicated as the original one. That is, in this section we have simply believed that there exists a black box method for reducing the initial number of opinions, and this method (for some reason) can not be used to reduce the initial number of opinions to a single one (otherwise we had solved our initial problem already). Consequently, the potential good news that we have presented in this section should be taken with some caution.

Discussion/Conclusion. In this chapter, we have analyzed **four modifications**

[O,K,n]	P_{bound}	$\mathbb{E}(p_e) = \frac{1}{O^{n-1}}$	Estimates of $P(D > 0)$
[2,1,2]	0.25	0.5	0.2885
[2,1,2]	0.375	0.5	0.4692
[2,1,2]	0.625	0.5	0.6944
[2,1,2]	0.75	0.5	0.6884
[3,1,2]	0.1667	0.3333	0.5941
[3,1,2]	0.25	0.3333	0.7102
[3,1,2]	0.4167	0.3333	0.7712
[3,1,2]	0.5	0.3333	0.7718
[4,1,2]	0.125	0.25	0.7244
[4,1,2]	0.1875	0.25	0.787
[4,1,2]	0.3125	0.25	0.8183
[4,1,2]	0.375	0.25	0.8163
[2,2,2]	0.25	0.5	0.3817
[2,2,2]	0.375	0.5	0.5276
[2,2,2]	0.625	0.5	0.7336
[2,2,2]	0.75	0.5	0.7347
[3,2,2]	0.1667	0.3333	0.596
[3,2,2]	0.25	0.3333	0.7177
[3,2,2]	0.4167	0.3333	0.783
[3,2,2]	0.5	0.3333	0.7867
[4,2,2]	0.125	0.25	0.6990
[4,2,2]	0.1875	0.25	0.7869
[4,2,2]	0.3125	0.25	0.8158
[4,2,2]	0.375	0.25	0.8177
[2,3,2]	0.25	0.5	0.4081
[2,3,2]	0.375	0.5	0.5364
[2,3,2]	0.625	0.5	0.7505
[2,3,2]	0.75	0.5	0.7464
[3,3,2]	0.1667	0.3333	0.5979
[3,3,2]	0.25	0.3333	0.7129
[3,3,2]	0.4167	0.3333	0.7857
[3,3,2]	0.5	0.3333	0.7874
[4,3,2]	0.125	0.25	0.6987
[4,3,2]	0.1875	0.25	0.7833
[4,3,2]	0.3125	0.25	0.817
[4,3,2]	0.375	0.25	0.8147

Table 4.3: Estimates of $P(D > 0)$ in which a bound on the probability of early agreements was set. The computation involved 20000 experiments for each profile.

of our original notion of model of deliberation (presented in Chapter 1). That is, we allowed for a **random selection** of the next speaker in debates. We experimented with a **sequential opinion reduction** too. We limited the **emergence of early agreements** in deliberations as well. Finally, we also tested the idea of **reducing the number of initial opinions (or generating cliques of them instead)** before a deliberation starts. The intention behind these modifications was always the same: to obtain a plausible alternative (to the original notion of *model of deliberation*) in which the opinion of the first speaker was not too strong. Among these alternatives, the *limitation on early agreements* seems to be the most promising one.

Chapter 5

Conclusion

In this dissertation there are four essential moments, and our intention is that each of them contributed to the same cause: the study of the significance of the order of speakers in models of deliberation. **In a first moment** (Chapter 1), we presented a mathematical structure that allows us to represent the dynamics of models of deliberation in a proper way. In that chapter, we also identified *opinion trees* and *opinion tree structures* as important notions for our project. With this description of a model of deliberation in place, **in a second moment** (Chapter 2), we defined and solved two key problems on opinion tree structures: the *overall anchoring problem* and the *probability of anchoring problem*. From the former we learnt about the influence of some parameters (e.g., the number of speakers, the number of opinions and the number of rounds) on the expected final outcome of a debate. From the latter we learnt that this influence might be notably beneficial for the first speaker. Naturally, with these results at hand, a justified reaction is to believe that *opinion trees* in which the opinion of the first speaker is not the most influential one are (sadly) rare mathematical objects (in the world of opinion trees). **In a third moment** (Chapter 3), we focused on the following situation: a single model of deliberation is given (as a *black box*), and we need to answer the question– *To which extent does this instance present order dependence/anchoring?* In this case, we did not approach this general question expecting to provide analytical results. Instead, the intention was to describe a general procedure (via simulations) for obtaining an estimation of the answer to our triggering question. The motivation behind this approach is that even after the negative results of Chapter 2, *we could still have models of deliberation that do not give a clear advantage to the first speaker*, and in these particular cases, we would like to have some test(s) that supported/rejected the claim "my model is not biased". **A fourth moment** (Chapter 4) was the stage for a refreshing model-exploration, and we investigated four natural modifications of our original notion of *model of deliberation*. Right from the beginning, in this investigation there was a declared hope: that perhaps at least one of these modifications could decrease the impact of the first speaker's opinion on the outcome of deliberations. Looking back, this exploration was both a pleasant exercise of trust (in our original notion of model) and a bittersweet experience (regarding the outcome of our exploration).

What could we do next? That is, assuming that our previous results are correct, *how could we react to that information?* By itself, this an interesting academic question, and if one is positive about the idea "having good deliberations is possible/needed", this question should be a pragmatic one as well. Next, we sketch some steps in which we appreciate a potential answer to the previous natural query.

State the results in their minimal alarming form. That is, the main results of this dissertation are **negative**. However, it would be great to keep two circumstances in mind. First, these results intend to say nothing (at least directly) regarding real-life scenarios. Instead, they speak about deliberative situations in the context of models (and these are theoretical objects). Second, (with the exception of four *model modifications* analyzed in Chapter 4) our results do not cover subfamilies (or alternatives) of diverse forms of deliberative situations. So, ours are global results, which were not proved true in restricted scenarios.

More exploration of individual models of deliberation. Regarding deliberative situations, our results have a global-scale nature. So, it might be perfectly plausible that the unwelcome *order-dependence* phenomenon behaved differently in local-scale scenarios (this is the idea to investigate). Consequently, an additional exploration of individual models of deliberation could be a constructive next step (this is how we investigate it). Basically, this step would require more experimenting with brand new dynamics of models of deliberation (close to what we did in Section 4.1 and Section 4.2). Alternatively, **instead** of searching for new dynamics, **one could** fix a single one **and** explore the impact of including extra information **regarding** the rationality of the individuals that deliberate (as in Hartmann and Rafiee Rad 2020). The intention behind these ideas is clear, we would like to find particular instances of models of deliberation in which the *order dependence* problematic was not present. Naturally, the outcome of this search might have different degrees of success.

Case 1: The results hold stable. Next, let us imagine for a moment that we conducted the steps described in the previous paragraph, and in most scenarios our results were stable (i.e., the first speaker had some advantage). Naturally, this would be unfortunate, and probably "*now*" would be the right moment for starting to consider more pragmatic questions. Accordingly, we describe two instances of such a questions here, and our hope is that they might still offer a real-life approach to the *order dependence* situation. **First**, *how strong is the impact of the order dependence?* Sadly, in this question we have already accepted that there might exist an unavoidable *order dependence*. However, there is a positive side as well; in cases in which the first speaker had only a negligible advantage, one could just disregard it (as we do "disregard" other minor events in critical decisions of our everyday life). It is worth noting that the previous question should be answered for each particular model that one considered as a "realistic one". Also, the precise meaning of the word *impact* is obviously the key to the question. In our view, this meaning should be related to a magnitude close to the expectation $\mathbb{E}(D)$ of D described in the *overall anchoring problem* (see Definition 8). That is, we would prefer to take the word *impact* as a term that is related to a sequence of experiments and not to a single one. However, this is a personal preference, and other interpretations for the same term might be more suitable in different circumstances. A **second** pragmatic question: assuming that *order dependence* is unavoidable, *what are the potential real-life implications?* Naturally, this question complements the previous one. We would like to know not only whether the *order dependence* issue is significant, but we want to know the precise real-life scenarios that might be affected by this problematic too. Clearly, "*a jump*" from *model results* to *real-life* implications requires empirical evidence as well (next paragraph addresses this point). However, assuming that neither *new models* nor *empirical evidence* provided us with good news, there are two principal kind of scenarios that might be negatively influenced by the *order dependence* matter (and both were disclosed in our introductory section). The first one consists of those deliberative situations that take place within frameworks of representative democracies. The second one comes from the private sector, and it includes decision-making situations related to corporations' boards, consulting agencies, etc. In both cases, a **new challenge** might be to **create, justify and keep trust in alternative mechanisms** for reaching consensus/agreement in generic instances of group decision-making (in the unfortunate case in which deliberations were intrinsically biased). Besides these two pragmatic questions, it is important to note that the more abstract one "*what are the implications of our results for deliberative democracy?*" would need to be addressed as well. That is, the term *deliberation* is at the core of this form of democracy, and this is problematic because the assumption "deliberative situations are unbiased" can not be taken for granted anymore.

Empirical evidence. Eventually, after properly squeezing our theoretical tools, we will need empirical evidence of the reported *order dependence*. A this stage, one could expect that a primary focus of an empirical study on deliberative situations would be to produce estimates of the expectation $\mathbb{E}(D)$ of D (related to the *overall anchoring problem*). A second focus of attention might be the estimation of $P(D > 0)$ (related to the *probability of anchoring problem*). In our view, the former aim could have a plausible/doable research path, while the second one might be quite challenging. Behind this conjecture, there are **two** credible reasons. **First**, the *overall anchoring problem* requires a succession of independent deliberative situations (which can be mimicked by a sequence of real-life experiments). **Second**, the *probability of anchoring problem* requires a one-shoot deliberative situation. In this case, a direct inference of estimates of probabilities based on a single deliberative experiment might be a difficult task. Besides these challenging ideas, there is an aspect regarding the context of experiments that is worth noting as well. It would be quite interesting to obtain (and contrast) empirical evidence of both kind of scenarios, those in which some experts deliberate and those in which non-experts need to decide on a given subject.

Case 2: The results do not hold stable. Next, let us imagine we are in the other branch of the stability-fork. That is, after some additional research (either model-based or empirical), we have established that in more *realistic/particular deliberative situations*, our general results do not hold anymore. Even in this convenient scenario, there were still **two** interesting questions that need to be answered. **First**, would it be possible to identify/characterize the exact families of deliberative situations in which the general results did not hold. **Second**, in the context of deliberative situations, is there any undesired by-product of the eradication of *order dependence*? Naturally, Section 4.3 offers a clear example of the kind of scenarios related to the first question. In that section, we realized that a limitation on early agreements (in deliberative situations) tends to decrease the strength of the first speaker's opinion. Moreover, in the same section, we also acquired a precise impression of how much do we need to limit those early agreements in order to see a positive impact on deliberative scenarios. Clearly, this "precise impression" was interesting because it provided us with a (non-formal) characterization of that class of deliberative situations.

As we have seen, there are many "ifs" in our idea of potential reactions to the results described in this dissertation. However, all those "ifs" are reasonably well founded and triggered by the existence of possible scenarios that might arise.

To conclude, in this dissertation we defend the idea that the structure of debates (as it was described here) does promote that the opinion of the first speaker is the (potentially) most influential one in a debate. Additionally, we understand that this phenomenon can be amplified by increasing the number of allowed opinions. Naturally, we believe that this information could be used to challenge the fairness of certain forms of debate. But, we also believe that this challenge is a positive one because it might encourage us all in the search for better models of deliberation.

Appendix A

Basic notions

This appendix contains basic notions that were required in this dissertation. Most of the definitions are direct extracts from the original sources, and they can be found in Durrett 2019 (for Measure and Probability Theory), Ng et al. 2011 or Kadane 2011 (for the notions on the Dirichlet distribution), Leobacher and Pillichshammer 2014 (for topics related to the Monte Carlo method), and West 2020 (for Graph Theory). Occasionally, a change of notation was needed (for the sake of uniformity); naturally, the meaning behind the notions remains untouched.

Notions of Measure and Probability Theory

σ-algebra. Given a set Ω, a σ-algebra $\mathcal{F}$ is a nonempty collection of subsets of Ω that satisfy: (i) if $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$ and (ii) if $A_i \in \mathcal{F}$ is a countable sequence of sets, then $\cup_i A_i \in \mathcal{F}$. Here and in what follows, countable means finite or countable finite.

Measurable Space. A measurable space is a pair $(\Omega, \mathcal{F})$ consisting of a set Ω and a σ-algebra $\mathcal{F}$ of subsets of Ω.

Measure. Given a measurable space $(\Omega, \mathcal{F})$, a measure is a non negative function $\mu : \mathcal{F} \to \mathbb{R}$ such that: (i) $\mu(A) \geq \mu(\emptyset) = 0$ for all $A \in \mathcal{F}$ and (ii) if $A_i \in \mathcal{F}$ is a countable sequence of disjoint sets, then $\mu(\cup_i A_i) = \sum_i \mu(A_i)$.

Probability Space. A probability space is a triple $(\Omega, \mathcal{F}, P)$, where Ω is a set of "outcomes", $\mathcal{F}$ is a set of "events" and $P : \mathcal{F} \to [0, 1]$ is a function that assigns probabilities to events. Formally, $\mathcal{F}$ is a σ-algebra, P is a measure on $(\Omega, \mathcal{F})$ and $P(\Omega) = 1$.

Borel sets. Given a set Ω and a collection $\mathcal{A}$ of subsets of Ω, there is a smallest σ-algebra containing $\mathcal{A}$. We call this, the σ-algebra generated by $\mathcal{A}$. Let $\mathbb{R}^d$ be the set of vectors $(x_1, x_2, \ldots, x_d)$ of real numbers, the Borel sets on $\mathbb{R}^d$ is the smallest σ-algebra containing the open sets. We denote it by $\mathcal{R}^d$, and when $d = 1$ we drop the superscript.

Random Variable. Given two measurable spaces $(\Omega, \mathcal{F})$ and $(S, \mathcal{S})$, a function $X : \Omega \to S$ is said to be a measurable map if the following condition holds: $X^{-1}(B) \equiv$

$\{\omega : X(w) \in B\} \in \mathcal{F}$ for all $B \in \mathcal{S}$. If $(S, \mathcal{S}) = (\mathbb{R}^d, \mathcal{R}^d)$ and $d > 1$, then X is called a random vector. In the particular case that $d = 1$, we say that X is a random variable. A simple, but useful, example of a random variable is the *indicator function* of a set $A \in \mathcal{F}$:

$$1_A(w) = \begin{cases} 1 & w \in A \\ \\ 0 & w \notin A \end{cases}$$

Expectation. Given a probability space $(\Omega, \mathcal{F}, P)$ and a random variable $X \geq 0$, then we define the expected value of X to be $\mathbb{E}(X) = \int X dP$. This expression is well defined, but it may be ∞. To reduce the general case ($X \geq 0$ or $X < 0$) to the non-negative case, let $x^+ = max\{x, 0\}$ be the positive part and let $x^- = max\{-x, 0\}$ be the negative part of x. We declare that $\mathbb{E}(X)$ exists and set $\mathbb{E}(X) = \mathbb{E}(X^+) - \mathbb{E}(X^-)$ whenever the subtraction makes sense (i.e., $\mathbb{E}(X^+) < \infty$ or $\mathbb{E}(X^-) < \infty$). Often, $\mathbb{E}(X)$ is called the mean of X and denoted by μ. Note, the integral symbol used above in the definition of expectation refers to the Lebesgue integral.

Variance. Given a probability space $(\Omega, \mathcal{F}, P)$ and a random variable X, if $\mathbb{E}(X^2) < \infty$, the variance $\mathbb{V}(X)$ of X is defined to be $\mathbb{E}((X - \mu)^2)$. To compute the variance the following formula is useful $\mathbb{V}(X) = \mathbb{E}(X^2) - \mu^2$.

Distribution. Given a probability space $(\Omega, \mathcal{F}, P)$, a random variable X induces a probability measure on $\mathbb{R}$ (called its distribution) by setting $\mu(A) = P(X \in A)$ for Borel sets A. The right-hand side can be written as $P(X^{-1}(A))$.

Chebyshev's Inequality. Suppose $\varphi : \mathbb{R} \to \mathbb{R}$ has $\varphi \geq 0$, let $A \in \mathcal{R}$ and let $i_A = \inf\{\varphi(y) : y \in A\}$. Then, (with the notation $\mathbb{E}(X; A) = \int_A X dP$) the following holds: $i_A P(X \in A) \leq \mathbb{E}(\varphi(X); X \in A) \leq \mathbb{E}(\varphi(X))$. In the particular case in which $\varphi(x) = x^2$ and $A = \{x : |x| \geq a\}$ we get $a^2 P(|X| \geq a) \leq \mathbb{E}(X^2)$. Some authors call the general case Markov's inequality, and use the name Chebyshev's inequality for the particular case.

Independence(Random variables). We say that the sequence of random variables $X_1, X_2, \ldots, X_n$ are independent if whenever $B_i \in \mathcal{R}$ for $i = 1, 2, \ldots, n$, we have $P(\cap_{i=1}^n \{X_i \in B_i\}) = \prod_{i=1}^n P(X_i \in B_i)$.

Law of Large Numbers(Strong). Let $X_1, X_2, \ldots, X_n$ be pairwise independent identically distributed random variables with $\mathbb{E}(|X_i|) < \infty$. Let $\mathbb{E}(X_i) = \mu$ and $S_n = X_1 + X_2, \cdots + X_n$. Then $S_n/n \to \mu$ almost surely as $n \to \infty$.

Dirichlet distribution

Dirichlet distribution. The Dirichlet distribution is a family of continuous multivariate probability distributions parameterized by a vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k)$ of positive reals. Next, we describe the probability density function (denoted as *pdf*), and the moments of the Dirichlet distribution.

Let S_k be the k-dimensional simplex, so

$$S_k = \{(x_1, x_2, \cdots, x_{k-1}) | x_i \geq 0; \sum_{i=1}^{k-1} x_i \leq 1\}$$

$$pdf := f(x_1, x_2, \ldots, x_{k-1}; \alpha) = \begin{cases} \frac{1}{B(\alpha)} \prod_{i=1}^{k} x_i^{\alpha_i - 1} & (x_1, x_2, \cdots, x_{k-1}) \in S_k \\ \\ 0 & \text{otherwise} \end{cases}$$

Note, in the previous expression the term x_k is a just an abbreviation for $1 - x_1 - x_2 - \cdots - x_{k-1}$. In addition, the multivariate beta function $B(\alpha)$ is defined as:

$$B(\alpha) = \frac{\prod_{i=1}^{k} \Gamma(\alpha_i)}{\Gamma(\sum_{i=1}^{k} \alpha_i)}$$

$$\Gamma(z) = \int_0^\infty e^{-x} x^{z-1} dx, \; \Re(z) > 0 \text{ (so that the integral converges absolutely).}$$

Here, $\Re(z)$ stands for the real part of z.

Particular case: $\Gamma(z) = (z-1)!\ , z \in \mathbb{Z}^+$

Further, the marginal distributions of the $x_i's$ are Beta distributions. That is, $x_i \sim Beta(\alpha_i, \bar{\alpha_i})$ with $\bar{\alpha_i} = (\sum_{m=1}^{k} \alpha_m) - \alpha_i$. This leads to the following cumulative distribution function, where B_I is the incomplete beta function :

$$\begin{aligned} CDF_{x_i}(x, \alpha_i, \bar{\alpha_i}) &= \frac{B_I(x; \alpha_i, \bar{\alpha_i})}{B(\alpha_i, \bar{\alpha_i})} \\ &= \frac{\int_0^x t^{(\alpha_i - 1)}(1-t)^{(\bar{\alpha_i} - 1)} dt}{B(\alpha_i, \bar{\alpha_i})} \end{aligned}$$

Regarding expectations, the general expression for the mixed moments is as follows:

$$\mathbb{E}(\prod_{t=1}^{k} x_t^{r_t}) = \frac{B(\alpha_1 + r_1, \alpha_2 + r_2, \ldots, \alpha_k + r_k)}{B(\alpha_1, \alpha_2, \ldots, \alpha_k)}$$

In the particular case in which $r_i = \ell$ is the only term different from zero:

$$\mathbb{E}(x_i^\ell) = \frac{(\alpha_i + \ell - 1)(\alpha_i + \ell - 2) \ldots (\alpha_i)}{(\sum_{j=1}^{k} \alpha_j + \ell - 1) \ldots (\sum_{j=1}^{k} \alpha_j)}$$

Consequently, if $\ell = 1$ or $\ell = 2$:

$$\mathbb{E}(x_i) = \frac{\alpha_i}{\sum_{j=1}^{k} \alpha_j}$$

$$\mathbb{E}(x_i^2) = \frac{(\alpha_i + 1)(\alpha_i)}{(\sum_{j=1}^{k} \alpha_j + 1)(\sum_{j=1}^{k} \alpha_j)}$$

Notions of Monte Carlo Integration

Intention. We aim at approximating the integral of a function $f : [0,1]^s \to \mathbb{R}$ by an equal weight quadrature rule of the form $\frac{1}{N} \sum_{n=0}^{N-1} f(x_n)$ where the quadrature points $\mathcal{P} = \{x_0, \ldots, x_{N-1}\}$ are of the form $[0,1)^s$.

We are interested in the integration error:

$$e(f, \mathcal{P}) := \int_{[0,1]^s} f(x) dx - \frac{1}{N} \sum_{n=0}^{N-1} f(x_n).$$

But, how should we choose the quadrature points? One idea is to choose realisations of N independent and uniformly distributed random variables $X_0, \ldots, X_{N-1}$ in $[0,1]^s$ and to check what we can expect for the resulting error. This means that we use

$$Q_{N,s}(f) := \frac{1}{N}\sum_{n=0}^{N-1} f(X_n)$$

as an estimator for the integral. Note that a measurable function $f : [0,1]^s \to \mathbb{R}$ can be considered as a random variable on the probability space $([0,1]^s, \mathcal{B}, \lambda_s)$ where $\mathcal{B}$ is the Borel σ-algebra on $[0,1]^s$ and λ_s the Lebesgue measure. Then the expectation of this random variable equals the integral we want to compute, i.e., $\mathbb{E}(f) = \int_{[0,1]^s} f(x)dx$. Using the linearity of the expected value we have

$$\mathbb{E}(Q_{N,s}(f)) = \frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}(f) = \mathbb{E}(f) = \int_{[0,1]^s} f(x)dx,$$

and hence $Q_{N,s}(f)$ is an unbiased estimator for the integral $\int_{[0,1]^s} f(x)dx$. The strong law of large numbers guarantees that

$$\mathbb{P}(\lim_{N\to\infty} Q_{N,s}(f) = \int_{[0,1]^s} f(x)dx) = 1,$$

where $\mathbb{P}$ is the probability on an arbitrary probability space supporting an independent sequence $(X_n)_{n\in\mathbb{N}_0}$ of random variables uniformly distributed on $[0,1]^s$.

The variance of f is given by $\mathbb{V}(f) := \int_{[0,1]^s}(f(x) - \int_{[0,1]^s} f(y)dy)^2 dx$. Since $X_0, \ldots, X_{N-1}$ are independent, we obtain from the Bienaymé formula the following result for the variance of the estimator $Q_{N,s}(f)$.

Variance of the estimator. Let $f \in L_2([0,1]^s)$. Then for any $N \in \mathbb{N}$ we have

$$\mathbb{V}(Q_{N,s}(f)) = \frac{\mathbb{V}(f)}{N}.$$

Note that

$$\mathbb{V}(Q_{N,s}(f)) = \mathbb{E}((Q_{N,s}(f) - \mathbb{E}(f))^2) = \mathbb{E}(e^2(f,\cdot)),$$

where $e^2(f,\cdot)$ is the error estimator

$$e(f,\cdot) := \int_{[0,1]^s} f(x)dx - \frac{1}{N}\sum_{n=0}^{N-1} f(X_n).$$

Hence it follows from the variance of the estimator that

$$\mathbb{E}(|e(f,\cdot)|) \le \sqrt{\mathbb{E}(e^2(f,\cdot))} = \frac{\sigma(f)}{\sqrt{N}},$$

where $\sigma(f) = \sqrt{\mathbb{V}(f)}$ denotes the standard deviation of f. This means that the absolute value of the integration error is, on average, bounded by $\sigma(f)/\sqrt{N}$. It is remarkable that the convergence rate of the expected integration error does not depend on the dimension s.

Advantages of the Monte Carlo method. First, it suffices that the integrands are quadratic integrable. Second, the convergence rate $O(\sqrt{N})$ is independent of the dimension s. This is a surprising fact, although it does not mean that the Monte Carlo method breaks the curse of dimensionality, because the standard deviation $\sigma(f)$ is in general not independent of s.

Disadvantages of the Monte Carlo method. First, the error bound is only "probabilistic", that is, in any one instance one cannot be sure of the integration error. However, further probabilistic information (bounds for N so that a certain confidence interval for the error bound is assured) is obtained from the central limit theorem. A second problem is that the generation of random samples is difficult. Third, for some applications the convergence rate of $O(\sqrt{N})$ is too slow. Fourth, the convergence rate $O(\sqrt{N})$ does not reflect some possible regularity of the integrand.

Notions of Graph Theory

Graph. A graph G is an ordered pair consisting of a vertex set $V(G)$ and an edge set $E(G)$, where each edge (element of $E(G)$) is a set of two vertices (elements of $V(G)$). The vertices of an edge are its endpoints. We write xy for an edge with endpoints x and y, and we say that x and y are neighbors. The order of a graph G is $|V(G)|$. A graph with order n is an n-vertex graph.

Path. A path with n vertices is a graph whose vertices can be named $v_1, v_2, \ldots, v_n$ so that the edges are $\{v_i v_{i+1} : 1 \leq i \leq n-1\}$. We write $\langle v_1, v_2, \ldots, v_n \rangle$ to specify a path having vertices $v_1, v_2, \ldots, v_n$ in order.

Cycle. A cycle with n vertices is a graph whose vertices can be named $v_1, v_2, \ldots, v_n$ so that the edge set is $\{v_i v_{i+1} : 1 \leq i \leq n-1\} \cup \{v_n v_1\}$. The length of a path or cycle is the number of edges.

Connected graph. A u, v-path is a path with first and last vertices u and v, called its endpoints. A graph G is connected if it contains a u, v-path for all $u, v \in V(G)$. If u and v are vertices in a connected graph G, their distance $d(u, v)$ is the length of the shortest u, v-path.

Tree. An acyclic graph is a graph with no cycles. A tree is a connected acyclic graph.

Rooted tree. A rooted tree has one vertex distinguished as a root. In a tree with root r, the neighbor of a vertex v on the path from v to r is the parent of v (i.e., the root has no parent), and the other neighbors of v are its children. A leaf in a rooted tree is a vertex with no children. If a vertex is not a leaf then it is internal. (Whenever the term "tree" appears in the content of this dissertation, it means rooted tree.)

Probabilistic tree. A weighted tree (T, f) consist of two elements, T is a tree and $f : E(T) \rightarrow \mathbb{R}$ is a mapping from the sets of edges of T onto the set of reals numbers. A probabilistic tree is a weighted tree in which: the weights of edges from every parent to its children add up to exactly one. This kind of trees is also known as *tree diagrams.*

Labeled tree. A triple (T, L, g) is a labeled tree if T is a tree, L is a finite set and $g : V(T) \rightarrow L$ is a mapping that labels each vertex of T with an element of L.

Weight of a leaf Given a probabilistic tree (T, f) with root r, the weight a leaf v in T is the multiplication of the weights of edges in the r, v-path.

Appendix B

Proofs of Statements

This appendix contains the proofs of theorems or lemmas that were presented in this dissertation.

Proof. (Theorem 1)
$\mathbb{E}(D) = D$ by Lemma 2.
$D = \sum_{i=1}^{\ell} L_i^{A,\Delta} p^i$ by definition of D and $L_i^{A,\Delta}$.
$D = \sum_{i=1}^{\ell} (-1)^{i+1} p^i$ by Lemma 3.
$D = \frac{p+(-p)^{\ell+1}}{p+1}$ by Lemma 4

■

Proof. (Lemma 2) We need to show that $\mathbb{E}(D) = D$. Because of the linearity of the expectation, the expression $\mathbb{E}(D)$ expands as the additions/subtractions of the expectation of terms in D. Now, if we assume that $\mathbb{E}(D) \neq D$, there should exist at least a term $\mathcal{T}$ in D, such that $\mathbb{E}(\mathcal{T}) \neq \mathcal{T}$. But, that can not occur: any such a term $\mathcal{T}$ is just the product of independent edge-probabilities (each of them with expectation equal to p). This means that $\mathbb{E}(\mathcal{T})$ is multiplicative, and it can be expanded as the multiplication of the expectation of the edge-probabilities in $\mathcal{T}$. This led to $\mathbb{E}(\mathcal{T}) = \mathcal{T}$, which was in contradiction with our assumption. ■

Proof. (Lemma 3) Next, we will need to count consensus vertices of an *opinion tree*, and that means we can ignore edge-probabilities for now. In a first step (the proof has two in total), we are interested in counting vertices L_i^A that: are at a distance $(i-1)$ from the root, and are still active (i.e., they are not consensus vertices). In this case, the idea to follow is: active vertices at the distance $(i-1)$ from the root are **exactly the result of** the branching of active vertices at a distance $(i-2)$ **minus** those vertices that resulted from the branching and became consensus vertices. That is, each of the L_{i-1}^A vertices produces $|O|^{n-1}$ vertices on the branching, and each time (exactly) one of them becomes a consensus vertex (see Example 3). This idea can be described as:

$$\begin{aligned} L_i^A &= L_{i-1}^A |O|^{n-1} - L_{i-1}^A && i > 1 \\ L_i^A &= 1 && i = 1 \end{aligned} \tag{B.1}$$

Consequently, the closed-form of the previous expression (see Appendix C.2 for details) is:

$$L_i^A = (|O|^{n-1} - 1)^{i-1} \qquad i \geq 1 \tag{B.2}$$

Using this result, in the second step we would need to count vertices $L_i^{A,x}$ that are active, at a distance $(i-1)$ from the root, and the current speaker has a specific opinion "x". In this case, the recursive idea is: the vertices that we want to count are **the result of** branching $L_{i-1}^A - L_{i-1}^{A,x}$ vertices (i.e., active with current opinion different from "x") **plus** the result of branching $L_{i-1}^{A,x}$ vertices (active with current opinion equal to "x"). Note, the reason for this distinction is that the number of vertices that results from the branching is different in each case. In the first one, a new opinion vertex will have **two opinions already set** (the opinion different from "x" required from the previous vertex **and** the opinion equal to "x" required in the new vertex). In the second case, each branching will result in (exactly) one new vertex less than in the first case (the one that will not be active because is a consensus vertex with all opinions equal to "x"). The previous idea can be described as:

$$\begin{aligned} L_i^{A,x} &= (L_{i-1}^A - L_{i-1}^{A,x}).|O|^{n-2} + L_{i-1}^{A,x}.(|O|^{n-2} - 1) && i > 1 \\ L_i^{A,x} &= 1 \qquad x = 0 \text{ (case of the first speaker)} && i = 1 \\ L_i^{A,x} &= 0 \qquad x = a \text{ (any opinion different from "0")} && i = 1 \end{aligned} \tag{B.3}$$

Before we present the close-form expressions for the above recurrence relations (note, there are two of them because of the two different initial conditions), let us adopt the following notation:

$$\begin{aligned} H &= |O|^{n-2} \\ Q &= |O|^{n-1} - 1 \end{aligned} \tag{B.4}$$

Now, the close-form solutions: the first equation corresponds to the count of consensus vertices **with** opinions equal to the one of the first speaker **and** distance i to the root of the opinion tree. Consequently, the second one corresponds to the count of consensus vertices **with** opinion different to the one of the first speaker **and** distance i to the root (see Appendix C.3 for details).

$$L_i^{A,0} = \frac{(-1)^i(HQ^i(-1)^i + Q(-Q + H - 1))}{Q(Q+1)} \qquad i \geq 1 \tag{B.5}$$

$$L_i^{A,a} = \frac{H(-1)^i((-1)^i Q^i + Q)}{Q(Q+1)} \qquad i \geq 1 \tag{B.6}$$

From these equations, we can immediately get that $L_i^{A,0} - L_i^{A,a} = L_i^{A,\Delta} = (-1)^{i+1}$. ■

Proof. (Lemma 4) We start with the second expression (in which ℓ is even). Then, the other case (in which ℓ is odd) is obtained by adding $p^{\ell+1}$ to the result of the $\ell + 1$ even case.

$p - p^2 + p^3 - p^4 + \cdots - p^\ell$
$= p(1-p) + p^3(1-p) + p^5(1-p) + \cdots + p^{\ell-1}(1-p)$
$= (1-p)(p + p^3 + p^5 + \cdots + p^{\ell-1})$
$= (1-p) \cdot \frac{p^{\ell+1} - p}{p^2 - 1}$

For the derivation of the last step we took $S_\ell = p + p^3 + p^5 + \cdots + p^{\ell-1}$ and $S_{\ell+2} = p+p^3+p^5+\cdots+p^{\ell+1}$. From here, the closed form of S_ℓ is obtained by solving the two equation system $S_{\ell+2} = S_\ell + p^{\ell+1}$ and $S_{\ell+2} = p^2 \cdot S_\ell + p$. Then, the proof continue as:

$= (1-p) \cdot \frac{p^{\ell+1}-p}{(p+1)(p-1)} \cdot \frac{-1}{-1}$

$= \frac{p-p^{\ell+1}}{p+1}$

Next, the case in which ℓ is odd follows from the previous expression as follows: we take the sum until an even $\hat{\ell} = \ell + 1$. Then, if we add to that sum a $p^{\hat{\ell}}$ term, what remains corresponds to the sum of the odd case: $\frac{p-p^{\hat{\ell}+1}}{p+1} + p^{\hat{\ell}} = \frac{p+p^{\hat{\ell}}}{p+1}$. Then, we obtain $\frac{p+p^{\ell+1}}{p+1}$.

■

Proof. (Theorem 5) Let us assume that the preconditions of Eq. 2.4 and Eq. 2.5 hold. That is: (O, K, n) is given, O can be arbitrary large, and the edge-probabilities for (O, K, n) will be selected in a single pick (keep in mind that uniformity and independence constraints are in place). Then, Lemma 10 and Lemma 11 can be applied. From the second one we get immediately that $P(S(\overrightarrow{0}) \leq B) \leq 1 - (\frac{1}{e})^{\frac{1}{N}}$. From the first one, if we take into account that O can be arbitrary large, $P(S(\overrightarrow{a}) \geq B) \leq \epsilon$ follows immediately as well.

■

Proof. (Lemma 6) Roughly speaking, the scheme of the proof is: first, we express $\mathbb{E}(S(\overrightarrow{a}))$ in a pleasant form. Then, we take the limit of the above expression, and show that it is equal to one. That is, $S(\overrightarrow{a}) = \sum_{i=1}^{\ell} E_i$, where E_i corresponds to the contribution of those "a-consensus" leaves that are at a distance i from the root (the number of these leaves was already denoted by $L_i^{A,a}$, see Eq. B.6). Moreover, because of the definition of $S(\overrightarrow{a})$, we have that $E_i = \sum_{j=1}^{L_i^{A,a}} p_{j,1}^i \cdot p_{j,2}^i \dots p_{j,i}^i$. In this summation (because an edge can participate in many paths to different consensus vertices) the subscripts and superscripts of edge-probabilities keep track of helpful descriptive information. Subscripts hold two indexes: the first is a reference to the particular consensus node; the second is the distance to the root of each particular edge-probability. Superscripts hold a single index, and it keeps track of the distance to the root at which the corresponding consensus node stands.

With $S(\overrightarrow{a})$ properly represented, the next step is to take its expectation:

$\mathbb{E}(S(\overrightarrow{a})) = \sum_{i=1}^{\ell} \sum_{j=1}^{L_i^{A,a}} \mathbb{E}(p_{j,1}^i \cdot p_{j,2}^i \dots p_{j,i}^i)$ due to the linearity of the expectation.

$= \sum_{i=1}^{\ell} \sum_{j=1}^{L_i^{A,a}} \mathbb{E}(p_{j,1}^i) \cdot \mathbb{E}(p_{j,2}^i) \dots \mathbb{E}(p_{j,i}^i)$ because the edge-probabilities under different parents are independent.

$= \sum_{i=1}^{\ell} \sum_{j=1}^{L_i^{A,a}} p^i$ by definition of expectation under an uniform Dirichlet distribution (see the Dirichlet distribution section of Basic Notions for more details).

$= \sum_{i=1}^{\ell} p^i \cdot L_i^{A,a}$.

$= \sum_{i=1}^{\ell} p^i \cdot (\frac{HQ^i + (-1)^i HQ}{Q(Q+1)})$ by characterization of $L_i^{A,a}$ (see Eq. B.6).

$= \frac{H \cdot p(-(-p)^l + (p \cdot Q)^l + p(-1+(-1+(-p)^l)Q + (p \cdot Q)^l))}{(1+p)(1+Q)(-1+p \cdot Q)}$ (see Appendix C.5 for details of the calculation)

Now, after replacing p, H, Q (see Theorem 1 and Eq. B.4 for details about their values), and taking the corresponding limit, we get the desired result (see Appendix C.6 for details of the calculation). ■

Proof. (Lemma 7) We start with $\mathbb{V}(S(\overrightarrow{a})) = \mathbb{E}(S(\overrightarrow{a})^2) - (\mathbb{E}(S(\overrightarrow{a})))^2$, which is the usual characterization of variance. Next, as we did in the previous lemma, we can unfold $S(\overrightarrow{a})$ as $S(\overrightarrow{a}) = \sum_{i=1}^{\ell} E_i$, where E_i corresponds to the contribution of those "a-consensus" leaves that are at a distance i of the root (the number of these leaves was already denoted by $L_i^{A,a}$, see Eq. B.6). Then we obtain:

$\mathbb{V}(S(\overrightarrow{a})) = \mathbb{E}((\sum_{i=1}^{\ell} E_i)^2) - (\mathbb{E}(S(\overrightarrow{a})))^2$

$\mathbb{V}(S(\overrightarrow{a})) = \mathbb{E}(\sum_{i=1}^{\ell}\sum_{k=1}^{\ell} E_i \cdot E_k) - (\mathbb{E}(S(\overrightarrow{a})))^2$ carrying out the multiplication.

$\mathbb{V}(S(\overrightarrow{a})) = \sum_{i=1}^{\ell}\sum_{k=1}^{\ell} \mathbb{E}(E_i \cdot E_k) - (\mathbb{E}(S(\overrightarrow{a})))^2$ by the linearity of the expectation.

$\mathbb{V}(S(\overrightarrow{a})) \leq \sum_{i=1}^{\ell}\sum_{k=1}^{\ell} L_i^{A,a} \cdot L_k^{A,a} \cdot (\frac{2}{(O^{n-1}+1)\cdot(O^{n-1})})^{\frac{i+k}{2}} - (\mathbb{E}(S(\overrightarrow{a})))^2$ by Lemma 8.

$\mathbb{V}(S(\overrightarrow{a})) \leq \sum_{i=1}^{\ell}\sum_{k=1}^{\ell} L_i^{A,a} \cdot L_k^{A,a} \cdot (\frac{2}{(O^{n-1}+1)\cdot(O^{n-1})})^{\frac{i+k}{2}}$ because $(\mathbb{E}(S(\overrightarrow{a})))^2$ is positive. Now, this leads us to the definition of $F(O)$:

$$F(O) \coloneqq \sum_{i=1}^{\ell}\sum_{k=1}^{\ell} L_i^{A,a} \cdot L_k^{A,a} \cdot (\frac{2}{(O^{n-1}+1)\cdot(O^{n-1})})^{\frac{i+k}{2}}$$

Moreover, this expression fulfill the limit requirement of the lemma as well (see Appendix C.7 for the close-form of $F(O)$, and Appendix C.8 for the details of the limit). ■

Proof. (Lemma 8) Before starting, let us recall that the probability contribution of "a-consensus" leaves (at distances i and k from the root of the opinion tree) can be expressed as $E_i = \sum_{j=1}^{L_i^{A,a}} p_{j,1}^i \cdot p_{j,2}^i \dots p_{j,i}^i$ and $E_k = \sum_{m=1}^{L_k^{A,a}} p_{m,1}^k \cdot p_{m,2}^k \dots p_{m,k}^k$. Now, in the multiplication $E_i \cdot E_k$ there are exactly $L_i^{A,a} \cdot L_k^{A,a}$ terms. That is, **if the expectation of each term was bounded by** $(\frac{2}{(O^{n-1}+1)\cdot(O^{n-1})})^{\frac{i+k}{2}}$ **we had the desired result (the linearity of the expectation is needed here as well)**. To see that this is indeed the case, let us consider the expression that results from the multiplication of the j term in E_i and the m term in E_k . The result of this multiplication has the form $p_{j,1}^i \cdot p_{j,2}^i \dots p_{j,i}^i \cdot p_{m,1}^k \cdot p_{m,2}^k \dots p_{m,k}^k$, which contains $i+k$ factors. Also, in the context of the previous term, the following holds: for each of the involved random variables, there can be at most one other random variable with the same parent node(i.e., we are working with root-to-leaf paths). Taken together, this fact and the independence assumption allow us to transform $\mathbb{E}(p_{j,1}^i \cdot p_{j,2}^i \dots p_{j,i}^i \cdot p_{m,1}^k \cdot p_{m,2}^k \dots p_{m,k}^k)$ into the product of the expectations of at least $\lceil \frac{(i+k)}{2} \rceil$ independent factors(each of them involving two random variables at most). For instance, if $i = k$ and $j = m$, the mentioned term looked as follows (note, this is just a punctual example):

$$\mathbb{E}(p_{j,1}^i \cdot p_{j,2}^i \dots p_{j,i}^i \cdot p_{m,1}^k \cdot p_{m,2}^k \dots p_{m,k}^k) = \mathbb{E}(p_{j,1}.p_{m,1}).\mathbb{E}(p_{j,2}.p_{m,2}) \dots \mathbb{E}(p_{j,i}.p_{m,k})$$

Naturally, in a different example (where the variables in some duos were independent) the expectation of a duo can be further split. Now, given that we are trying to obtain an upper-bound on the product of expectations, we need to find the best possible arrangement for this product. As we will see, Lemma 9 offers that exactly; it tells that keeping as many duos as possible (and assuming that both elements of the duos are identical to each other) is the best way to proceed. For the last step, we just need to appreciate that in the original term $p_{j,1}^i \cdot p_{j,2}^i \dots p_{j,i}^i \cdot p_{m,1}^k \cdot p_{m,2}^k \dots p_{m,k}^k$ there can be at most $\lfloor \frac{(i+k)}{2} \rfloor$ duos of random variables. Taken together, this fact, the linearity of expectation, and the monotonicity of the square root function for the case in which $i + k$ is odd, give us the required bound. ■

Proof. (Lemma 9) We start with $\mathbb{E}(\prod_{t=1}^{\mu} p_t^{r_t}) = \frac{B(\alpha_1+r_1,\alpha_2+r_2,\ldots,\alpha_\mu+r_\mu)}{B(\alpha_1,\alpha_2,\ldots,\alpha_\mu)}$, which is the general expression for the mixed moments of a random vector $(p_1, p_2, \cdots, p_\mu)$ that follows a Dirichlet distribution (see the Dirichlet distribution section of Basic Notions). Here we use μ as a short name of O^{n-1}, which **is** the number of outgoing edge-probabilities **in** any internal vertex **of** an *opinion tree structure*. Next, we use this expression to derive the expectations terms that we are interested in (let us keep in mind that because of the uniformity assumption, the alphas are all equal to one).

$$\mathbb{E}(p_i) = \frac{B(\alpha_1+0,\ldots,\alpha_i+1,\ldots,\alpha_\mu+0)}{B(\alpha_1,\alpha_2,\ldots,\alpha_\mu)} = \frac{B(1,\ldots,2,\ldots,1)}{B(1,\ldots,1,\ldots,1)} = \frac{\frac{\Gamma(1)\ldots\Gamma(2)\ldots\Gamma(1)}{\Gamma(\mu+1)}}{\frac{\Gamma(1)\ldots\Gamma(1)\ldots\Gamma(1)}{\Gamma(\mu)}} = \frac{1}{\mu} = \frac{1}{O^{n-1}}$$

$$\mathbb{E}(p_i^2) = \frac{B(\alpha_1+0,\ldots,\alpha_i+2,\ldots,\alpha_\mu+0)}{B(\alpha_1,\alpha_2,\ldots,\alpha_\mu)} = \frac{B(1,\ldots,3,\ldots,1)}{B(1,\ldots,1,\ldots,1)} = \frac{\frac{\Gamma(1)\ldots\Gamma(3)\ldots\Gamma(1)}{\Gamma(\mu+2)}}{\frac{\Gamma(1)\ldots\Gamma(1)\ldots\Gamma(1)}{\Gamma(\mu)}} = \frac{2}{(O^{n-1}+1)(O^{n-1})}$$

$$\mathbb{E}(p_i \cdot p_j) = \frac{B(\alpha_1+0,\ldots,\alpha_i+1,\ldots,\alpha_j+1,\ldots,\alpha_\mu+0)}{B(\alpha_1,\alpha_2,\ldots,\alpha_\mu)} = \frac{B(1,\ldots,2,\ldots,2,\ldots,1)}{B(1,\ldots,1,\ldots,1)} = \frac{1}{(O^{n-1}+1)(O^{n-1})}$$

From these we can immediately see that $\mathbb{E}(p_i^2) > \mathbb{E}(p_i \cdot p_j)$, which is the second inequality that we wanted to show. In the case of the inequality $\mathbb{E}(p_i^2) > (\mathbb{E}(p_i))^2$, first we can prove that for any natural $y \geq 2$, it holds that $\frac{2}{(y+1)y} > \frac{1}{y^2}$. Which is equivalent to show that $\frac{2}{(y+1)} > \frac{1}{y}$. Naturally, if that was so, we could then make $y = O^{n-1}$ and obtain the desired result. The proof of the intermediate step can be done as follows:

$$2y > y + 1$$
$$\frac{2y}{(y+1)y} > \frac{y+1}{(y+1)y}$$
$$\frac{2}{(y+1)} > \frac{1}{y}$$

■

Proof. (Lemma 10) In this proof we use a simple form of the Chebyshev's Inequality (see the Probability Theory section of Basic Notions). Broadly said, for a random variable X, this inequality bounds by $\frac{1}{k^2}$ the portion of distribution that can be k or more standard deviations away from the mean of X. That is,

$$P(|X - \mu| \geq k \cdot \sigma) \leq \frac{1}{k^2}$$

Next, we can apply this inequality to our problem by taking $X = S(\overrightarrow{a})$ and $k \cdot \sigma = \frac{\mathbb{E}(D)}{N}$. Consequently, **what we need to show is**:

$$\frac{1}{k^2} \leq \frac{N \cdot \sqrt{2 \cdot ((3 + 2\sqrt{2}) \cdot (2^{\frac{l+1}{2}} - 2)^2)}}{O}$$

In order to do this, **we aim to prove that** $k > 1$ and:

$$\frac{1}{k} \leq \frac{N \cdot \sqrt{2 \cdot (3 + 2\sqrt{2}) \cdot (2^{\frac{l+1}{2}} - 2)^2}}{O}$$

So, we begin with **our above assumption**: $k = \frac{\mathbb{E}(D)}{N \cdot \sigma}$.

Consequently, from this assumption and Eq. 2.2, **we get** that $k = \frac{p(1+o(1))}{N \cdot \sigma}$. Here we used that $\mathbb{E}(D)$ and p are asymptotically equal as $O \to \infty$, which was expressed in the standard Little-o notation (for a description of the asymptotic equality relation, see Shabunin 2022, **and** for a definition and examples of the Little-o notation, see Cormen et al. 2022, p. 60). In the scope of this proof, from all the elements of $o(1)$, which help to approximate $\mathbb{E}(D)$ as $p(1 + o(1))$, we will work with those that take positive values only. That is, strictly speaking we will be working with a subset of $o(1)$ (i.e., we follow the definition of Little-o provided in Cormen et al. 2022, p. 60). For instance, our restriction allows $\frac{1}{O}$ in $o(1)$ as $O \to \infty$, but forbids $-\frac{1}{O}$. Naturally, the reason for this restriction will soon become apparent. Next, **we obtain**:

$$k \geq \frac{p(1 + o(1))}{N \cdot \sqrt{\frac{(3+2\sqrt{2})\cdot(2^{\frac{l+1}{2}}-2)^2}{O^{2n}}(1 + o(1))}}$$

In this inequality we do not have σ anymore. Instead, we have an upper bound on it. Let us recall that by definition $\sigma = \sqrt{\mathbb{V}(S(\overrightarrow{a}))}$. Then, from Lemma 7 **we know that** $\mathbb{V}(S(\overrightarrow{a})) \leq F(O)$ **and that** $F(O)$ is asymptotically equal to $\frac{(3+2\sqrt{2})\cdot(2^{\frac{l+1}{2}}-2)^2}{O^{2n}}$ as $O \to \infty$. That is, $F(O) = \frac{(3+2\sqrt{2})\cdot(2^{\frac{l+1}{2}}-2)^2}{O^{2n}}(1 + o(1))$. Together, these properties and the monotonicity of the square root function justify the inequality. Next, by decreasing the numerator we keep the inequality.

$$k \geq \frac{p}{N \cdot \sqrt{\frac{(3+2\sqrt{2})\cdot(2^{\frac{l+1}{2}}-2)^2}{O^{2n}}(1 + o(1))}}$$

Next, we increase the denominator and keep the inequality:

$$k \geq \frac{p}{N \cdot \sqrt{2 \cdot \frac{(3+2\sqrt{2})\cdot(2^{\frac{l+1}{2}}-2)^2}{O^{2n}}}}$$

Note, the bound that we just used is $1 \geq o(1)$. This is a valid bound because we know that (by definition of the Little-o relation) for a function $f(O) \in o(1)$, it holds that: for any positive constant ϵ there exist a constant O_ϵ such that $|f(O)| \leq \epsilon$ for all $O \geq O_\epsilon$. Then, if we take $\epsilon = 1$, we can be sure that there exist an O_1 such that $|f(O)| \leq 1$ for all $O \geq O_1$. Consequently, the previous bound is correct as long as we mention that it holds "for a sufficiently large O", which is something that we have already assumed.

Next, we take the term O^{2n} out of the square root:

$$k \geq \frac{p}{\frac{N}{O^n} \cdot \sqrt{2 \cdot (3 + 2\sqrt{2}) \cdot (2^{\frac{l+1}{2}} - 2)^2}}$$

Now, after substituting p by its value (i.e., $p = \frac{1}{O^{n-1}}$):

$$k \geq \frac{\frac{1}{O^{n-1}}}{\frac{N}{O^n} \cdot \sqrt{2 \cdot (3 + 2\sqrt{2}) \cdot (2^{\frac{l+1}{2}} - 2)^2}}$$

After reducing the O^{n-1} similar terms:

$$k \geq \frac{O}{N \cdot \sqrt{2 \cdot (3+2\sqrt{2}) \cdot (2^{\frac{l+1}{2}} - 2)^2}}$$

Which, under the large O assumption leads us to $k > 1$ and:

$$\frac{1}{k} \leq \frac{N \cdot \sqrt{2 \cdot (3+2\sqrt{2}) \cdot (2^{\frac{l+1}{2}} - 2)^2}}{O}$$

■

Proof. (Lemma 11) We start with the marginal distribution of X_i (see the Dirichlet distribution section of Basic Notions).

$CDF_{X_i}(x, \alpha_i, \bar{\alpha_i}) = \frac{B_I(x;\alpha_i,\bar{\alpha_i})}{B(\alpha_i,\bar{\alpha_i})}$, where B_I is the incomplete beta function, and B is the beta function(used as a normalizing factor).

$= \frac{\int_0^x t^{(\alpha_i-1)}(1-t)^{(\bar{\alpha_i}-1)}dt}{B(\alpha_i,\bar{\alpha_i})}$, replacing B_I by its definition.

$= \frac{\int_0^x (1-t)^{(\bar{\alpha_i}-1)}dt}{B(1,\bar{\alpha_i})}$, because $(\alpha_i - 1)$ is zero.

$= \frac{\int_0^x (1-t)^{(\bar{\alpha_i}-1)}dt}{\frac{\Gamma(1)\cdot\Gamma(\bar{\alpha_i})}{\Gamma(1+\alpha_i)}}$, after replacing B by its definition.

$= \bar{\alpha_i} \cdot \int_0^x (1-t)^{(\bar{\alpha_i}-1)}dt$, after simplifying the denominator.

$= \bar{\alpha_i} \cdot \Big|_0^x - \frac{(1-t)^{\bar{\alpha_i}}}{\bar{\alpha_i}}$, after calculating the integral.

$= \Big|_0^x - (1-t)^{\bar{\alpha_i}}$, after simplifying the $\bar{\alpha_i}$ terms.

$= -(1-x)^{\bar{\alpha_i}} - (-(1-0)^{\bar{\alpha_i}})$, after evaluating the previous expression.

$= -(1-x)^{\bar{\alpha_i}} + 1$

$= 1 - (1 - (\mathbb{E}(S(\overrightarrow{a})) + \frac{\mathbb{E}(D)}{N}))^{(O^{(n-1)}-1)}$, after substituting x and $\bar{\alpha_i}$ with their respective values.

Next, we can take the limit of the previous expression when $O \to \infty$. Note, from Eq. 2.2 and Lemma 6, we know that $\lim_{O\to\infty} \frac{\mathbb{E}(D)}{\frac{1}{O^{(n-1)}}} = 1$ and $\lim_{O\to\infty} \frac{\mathbb{E}(S(\overrightarrow{a}))}{\frac{\ell-1}{O^n}} = 1$. That is, we know that $\mathbb{E}(S(\overrightarrow{a})) = o(\mathbb{E}(D))$, and this allows us to let the term $\mathbb{E}(S(\overrightarrow{a})$ out of the above expression in the calculation of the limit. Then, we can proceed as follows (see Appendix C.9 for the details of the calculation of the limit):

$$\lim_{O\to\infty} 1 - (1 - (\mathbb{E}(S(\overrightarrow{a})) + \frac{\mathbb{E}(D)}{N}))^{(O^{(n-1)}-1)}$$

$$= \lim_{O\to\infty} 1 - (1 - \frac{\mathbb{E}(D)}{N})^{(O^{(n-1)}-1)}$$

$$= 1 - (\frac{1}{e})^{\frac{1}{N}}$$

■

Appendix C

Source Code: Mathematica

This appendix contains code snippets in the Wolfram Language. The intention is that they serve as witnesses for minor analytical statements (limits, sums, etc.) that were used (without a proof) in the dissertation.

```
(*E(D)*)
(p+Power[-p,l+1])/(p+1)

(*Previous expression expanded: taking p== 1/Power[O,n-1]*)
(*Additionally, we take its division by 1/Power[O,n-1]*)

(((1/Power[O,n-1])+Power[-(1/Power[O,n-1]),l+1])/((1/Power[O,n-1])+1)) / (1/Power[O,n-1])

(*Limit E(D)/p*)
Limit[(((1/Power[O,n-1])+Power[-(1/Power[O,n-1]),l+1])/((1/Power[O,n-1])+1)) / (1/Power[O,n-1]),O->Infinity, Assumptions -> l >=n && n>=2 && l \[Element] Integers]
```

Listing C.1: $\lim_{O\to\infty} \frac{\mathbb{E}(D)}{p} = 1$

```
(*Solves the recurrence that counts the number of actives vertices at a distance i-1 of the root*)
RSolve[{a[i]==(a[i-1]*Power[O,n-1])-a[i-1],a[1]==1},a[i],i]
```

Listing C.2: Closed-Form Solution of L_i^A.

```
(*Solves the recurrence that counts the number of actives vertices that are: at a distance i-1 of the root, and with an opinion that is equal to the one of the first speaker*)

RSolve[{a[i]==(Power[Q,i-2]-a[i-1])*H+a[i-1]*(H-1),a[1]==1},a[i],i]

(*Solves the recurrence that counts the number of actives vertices that are: at a distance i-1 of the root, and  with an opinion that is different from that of the first speaker*)

RSolve[{a[i]==(Power[Q,i-2]-a[i-1])*H+a[i-1]*(H-1),a[1]==0},a[i],i]
```

Listing C.3: Close-Form Solution of $L_i^{A,x}$.

```
(*Expectation expression*)
f[O_] :=(Power[1/O,n-1]+Power[-1*Power[1/O,n-1],(n*K)+1])/(Power[1/O,n-1]+1)

(*Derivative of the Expectation*)
f'[O]

(*Simplify the previous result*)
Simplify[%]
```

Listing C.4: Derivative of $\mathbb{E}(D)$ wrt. O

```
(*Expectation of S(a)*)
Sum[(p^i*(H*Q^i + (-1)^i*H*Q))/(Q*(Q + 1)), {i, 1, l}]

(*Close form: this is the result of executing the previous line*)
(H p (-(-p)^l+(p Q)^l+p (-1+(-1+(-p)^l) Q+(p Q)^l)))/((1+p) (1+Q) (-1+p Q))


(*Previous expression expanded: that is, taking p== 1/Power[O,n-1], H==Power[O,n-2] and Q == Power[O,n-1]-1*)
(Power[O,n-2] (1/Power[O,n-1]) (-(-(1/Power[O,n-1]))^l+((1/Power[O,n-1]) (Power[O,n-1]-1))^l+(1/Power[O,n-1]) (-1+(-1+(-(1/Power[O,n-1]))^l) (Power[O,n-1]-1)+((1/Power[O,n-1]) (Power[O,n-1]-1))^l)))/((1+(1/Power[O,n-1])) (1+(Power[O,n-1]-1)) (-1+(1/Power[O,n-1]) (Power[O,n-1]-1)))
```

Listing C.5: Expectation of S(a) and its close-form.

```
(*Limit of E(S(a))/((l-1)/Power[O,n-1]) when O  goes to infinity*)
Limit[((Power[O,n-2] (1/Power[O,n-1]) (-(-(1/Power[O,n-1]))^l+((1/Power[O,n-1]) (Power[O,n-1]-1))^l+(1/Power[O,n-1]) (-1+(-1+(-(1/Power[O,n-1]))^l) (Power[O,n-1]-1)+((1/Power[O,n-1]) (Power[O,n-1]-1))^l)))/((1+(1/Power[O,n-1])) (1+(Power[O,n-1]-1)) (-1+(1/Power[O,n-1]) (Power[O,n-1]-1))))/((l-1)/Power[O,n]) ,O->Infinity, Assumptions -> l >=n&& n >=2 && l \[Element] Integers]

(*Taking the limit after the change of variable: x== Power[O,n-1]*)
Limit[(-(-(1/x))^l+((1/x) (-1+(x)))^l+(1/x) (-1+((1/x) (-1+(x)))^l+(-1+(x)) (-1+(-(1/x))^l)))/((l-1) (1+(1/x)) (-1+(1/x) (-1+(x)))),x->\[Infinity],Assumptions->l>=2&&l \[Element] Integers]
```

Listing C.6: Expectation of S(a) and its asymptotic behavior when $O \to \infty$.

```
(*Upper bound of V(S(a)): term with double sum)*)
Sum[((H*Q^i + (-1)^i*H*Q)/(Q*(Q + 1)))*
   ((H*Q^k + (-1)^k*H*Q)/(Q*(Q + 1)))*
   (2/((Q + 2)*(Q + 1)))^((i + k)/2), {i, 1, l}, {k, 1, l}]


(*Close form: this is the result of executing the previous line*)
(H^2 (2 + (-1)^l 2^((3 + l)/2) (1/(2 + 3 Q + Q^2))^((1 + l)/2) + (-1)^
     l 2^((1 + l)/2) Q^2 (1/(2 + 3 Q + Q^2))^((1 + l)/2) -
    3 2^((1 + l)/2) Q^(1 + l) (1/(2 + 3 Q + Q^2))^((1 + l)/2) -
```

```
     2^((1 + 1)/2) Q^(2 + 1) (1/(2 + 3 Q + Q^2))^((1 + 1)/2) -
     2^(1 + 1/2) Q^1 (1/(2 + 3 Q + Q^2))^(
      1/2) (1 + Sqrt[2] Sqrt[1/(2 + 3 Q + Q^2)]) +
     Q (2 + (-1)^(1 + 1) 2^(1 + 1/2) (1/(2 + 3 Q + Q^2))^(1/2) +
        3 (-1)^1 2^((1 + 1)/2) (1/(2 + 3 Q + Q^2))^((1 + 1)/
         2)))^2)/((1 + Q)^4 (2 + Q)^2 (1 +
     Sqrt[2] Sqrt[1/(2 + 3 Q + Q^2)])^2 (-1 +
     Sqrt[2] Q Sqrt[1/(2 + 3 Q + Q^2)])^2)



(*Previous expression expanded: that is, taking H==(Power[O,n-2]) and Q == (Power[O,n-1]-1) *)
((Power[O,n-2])^2 (2 + (-1)^1 2^((3 + 1)/2) (1/(2 + 3 (Power[O,n-1]-1) + (Power[O,n-1]-1)^2))^((1 + 1)/2) + (-1)^
      1 2^((1 + 1)/2) (Power[O,n-1]-1)^2 (1/(2 + 3 (Power[O,n-1]-1) + (Power[O,n-1]-1)^2))^((1 + 1)/2) -
     3 2^((1 + 1)/2) (Power[O,n-1]-1)^(1 + 1) (1/(2 + 3 (Power[O,n-1]-1) + (Power[O,n-1]-1)^2))^((1 + 1)/2) -
     2^((1 + 1)/2) (Power[O,n-1]-1)^(2 + 1) (1/(2 + 3 (Power[O,n-1]-1) + (Power[O,n-1]-1)^2))^((1 + 1)/2) -
     2^(1 + 1/2) (Power[O,n-1]-1)^1 (1/(2 + 3 (Power[O,n-1]-1) + (Power[O,n-1]-1)^2))^(
      1/2) (1 + Sqrt[2] Sqrt[1/(2 + 3 (Power[O,n-1]-1) + (Power[O,n-1]-1)^2)]) +
     (Power[O,n-1]-1) (2 + (-1)^(1 + 1) 2^(1 + 1/2) (1/(2 + 3 (Power[O,n-1]-1) + (Power[O,n-1]-1)^2))^(1/2) +
        3 (-1)^1 2^((1 + 1)/2) (1/(2 + 3 (Power[O,n-1]-1) + (Power[O,n-1]-1)^2))^((1 + 1)/
         2)))^2)/((1 + (Power[O,n-1]-1))^4 (2 + (Power[O,n-1]-1))^2 (1 +
     Sqrt[2] Sqrt[1/(2 + 3 (Power[O,n-1]-1) + (Power[O,n-1]-1)^2)])^2 (-1 +
     Sqrt[2] (Power[O,n-1]-1) Sqrt[1/(2 + 3 (Power[O,n-1]-1) + (Power[O,n-1]-1)^2)])^2)
```

Listing C.7: Upper bound on the variance of S(a) and its close-form.

```
(*Asymptotic behavior: F(O) when O goes to infinity*)
Limit[(((Power[O,n-2])^2 (2 + (-1)^1 2^((3 + 1)/2) (1/(2 + 3 (Power[O,n-1]-1) + (Power[O,n-1]-1)^2))^((1 + 1)/2) + (-1)^
      1 2^((1 + 1)/2) (Power[O,n-1]-1)^2 (1/(2 + 3 (Power[O,n-1]-1) + (Power[O,n-1]-1)^2))^((1 + 1)/2) -
     3 2^((1 + 1)/2) (Power[O,n-1]-1)^(1 + 1) (1/(2 + 3 (Power[O,n-1]-1) + (Power[O,n-1]-1)^2))^((1 + 1)/2) -
     2^((1 + 1)/2) (Power[O,n-1]-1)^(2 + 1) (1/(2 + 3 (Power[O,n-1]-1) + (Power[O,n-1]-1)^2))^((1 + 1)/2) -
     2^(1 + 1/2) (Power[O,n-1]-1)^1 (1/(2 + 3 (Power[O,n-1]-1) + (Power[O,n-1]-1)^2))^(
      1/2) (1 + Sqrt[2] Sqrt[1/(2 + 3 (Power[O,n-1]-1) + (Power[O,n-1]-1)^2)]) +
     (Power[O,n-1]-1) (2 + (-1)^(1 + 1) 2^(1 + 1/2) (1/(2 + 3 (Power[O,n-1]-1) + (Power[O,n-1]-1)^2))^(1/2) +
        3 (-1)^1 2^((1 + 1)/2) (1/(2 + 3 (Power[O,n-1]-1) + (Power[O,n-1]-1)^2))^((1 + 1)/
         2)))^2)/((1 + (Power[O,n-1]-1))^4 (2 + (Power[O,n-1]-1))^2 (1 +
```

```
     Sqrt[2] Sqrt[1/(2 + 3 (Power[O,n-1]-1) + (Power[O,n-1]-1)^2)])
    ^2 (-1 +
     Sqrt[2] (Power[O,n-1]-1) Sqrt[1/(2 + 3 (Power[O,n-1]-1) + (
    Power[O,n-1]-1)^2)])^2))/(((3 + 2 Sqrt[2]) (-2 + 2^((1 + l)/2))
    ^2 )/Power[O,2*n]),O->Infinity, Assumptions -> l >=n&& n>=2 &&
    l \[Element] Integers]
```

Listing C.8: Asymptotic behavior of $F(O)$ when $O \to \infty$.

```
(*CDF of X_i when O is large and x=E(D)/N*)
Limit[1-Power[(1-((((1/Power[O,n-1])+Power[-(1/Power[O,n-1]),l+1])
    /((1/Power[O,n-1])+1))/N)),(Power[O,(n-1)]-1)], O->Infinity,
    Assumptions -> l >=n&& n>=2 && l \[Element] Integers]
```

Listing C.9: Dirichlet marginal distribution: $O \to \infty$ and $x = \mathbb{E}(D)$.

Appendix D

Source Code: Simulations

This appendix contains the python code that was used for the experiments with opinion trees structures.

```
import numpy as np
from decimal import *

###Calculate Probabilities in tree structure(given tree structure and Opinions)###
def evalProbabilities(treeStructure, opinion):
    sumProbabilities= 0

    if(isLeaf(treeStructure)):
        if(filterAllAreOne(treeStructure[0],[opinion])):
            return treeStructure[2]
        else:
            return 0

    else:

       for tree in treeStructure[1]:

           sumProbabilities =  sumProbabilities + evalProbabilities(tree,opinion)

       #The term (treeStructure[2])* is the reason why the probability on the root is
       #set to one in the function generateTreeStructure.
       return (treeStructure[2])*sumProbabilities
###########Generate a tree structure###################
#This function creates the root of the tree (i.e., the opinionSpeaker
#variable describes the important information of the inital
#debate conditions).Also, treeStructure contains subtrees, their
#roots describe possible group opinions after the first speaker spoke.
def generateTreeStructure(O,n,K,opinionSpeaker,typeOfStructure):
  treeStructure=[]

  #Creates all possible permutations.
  allPermutations = fullPermutation(n,O)

  #The variable typeOfStructure allows to call the particular function that will
  #build the structure.Naturally, it will be
```

```
  #different depending on the type of experiment.
  treeStructure= typeOfStructure(O,n,K,opinionSpeaker,0,0,allPermutations,K*n-1)


  #The symbol * does not play any particular role. It is just a mark to know that
  #this is the first speaker's initial opinion.
  toReturn= [str(opinionSpeaker)+'*',treeStructure,1]
  assignProbabilitiesToTreeStructure(toReturn,0,n)
  #check that the given structure is correct.
  #print(checkAgreement(toReturn,0,n,opinionSpeaker,K*n), checkSumProbabilities(toReturn,0))
  #print (toReturn)
  return toReturn

#This is the actual recursive function that generates the tree.
#Note: distanceToComplete is an edge based distance that starts in the vertex
#that is to be built.
def generateTreeStructure1(O,n,K,opinionSpeaker,distanceToRoot,positionSpeaker,allPermutations, distanceToComplete):

      treeStructure=[]
      if(distanceToComplete==0):

          #This filter gets all permutations of possible group opinions that
          # agree with the opinion of the previous speaker at her position.
          filteredPermutations  = filterAll(allPermutations,[filterAtPosition],[[positionSpeaker,opinionSpeaker]])
          for permutation in filteredPermutations:

                #The 'd' just means that the node was bounded by the distance.
                #Note, agreement nodes can be marked with 'd'(e.g., whenever it
                #happens that the node is bounded by distance and it is
                #and agreement node too). Also, that can be checked by
                #searching the 'd' term in this script.
                treeStructure.append([permutation,'d',None])

      else:

          #Note: this line is in the other part of the condition too.
          #It is not outside of the condition for the sake of clearity.
          filteredPermutations  = filterAll(allPermutations,[filterAtPosition],[[positionSpeaker,opinionSpeaker]])
          #print(filteredPermutations)
          #exit()
          for permutation in filteredPermutations:

                #currentTreeTemp will remain 'a' if the permutation
                #was a consensus node.
                currentTreeTemp='a'

                #Note: in the next line one could probably use
                #the driver as well-->
                #filterAll(allPermutations,[filterNotAllAreOne],[[opinionSpeaker]])
                #instead of using a direct call.
                if(filterNotAllAreOne(permutation,[opinionSpeaker])):

                   #Note: permutation[(distanceToRoot)%n] should be the same
```

```
                   #as opinionSpeaker because permutation is obtained after
                   #filtering with filterAtPosition
                   currentTreeTemp =  generateTreeStructure1(O,n,K,permutation[(distanceToRoot+1)%n],distanceToRoot+1,(distanceToRoot+1)%n,allPermutations,distanceToComplete-1)
                #currentTreeTemp=='a' if this was an agreemment node with
                #the opinionSpeaker opinion.
                treeStructure.append([permutation,currentTreeTemp,None])

      return treeStructure
#######
#The following two functions will be used for the experiment in which the
#position of the next speaker is selected at random(uniformaly).

#This is the actual recursive function that generates the tree.
#Note: distanceToComplete is an edge based distance that starts
#in the vertex that is to be built.
#Note: This function describes the scenario where the position of
#the next speaker is selected at random.
def generateTreeStructure1_random(O,n,K,opinionSpeaker,distanceToRoot,positionSpeaker,allPermutations, distanceToComplete):

      treeStructure=[]
      if(distanceToComplete==0):

          #This filter gets all permutations of possible group opinions that agree
          #with the opinion of the previous speaker at her position.
          filteredPermutations  = filterAll(allPermutations,[filterAtPosition],[[positionSpeaker,opinionSpeaker]])
          for permutation in filteredPermutations:

                #The 'd' just means that the node was bounded by the distance.
                #Note, agreement nodes can be marked with 'd'(e.g., whenever
                #it happens that the node is bounded by distance and it is an
                #agreement node too). Also, that can be checked by searching
                #the 'd' term in this script.
                treeStructure.append([permutation,'d',None])

      else:

          #Note: this line is in the other part of the condition too. It is not
          #outside of the condition for the sake of clearity.
          filteredPermutations  = filterAll(allPermutations,[filterAtPosition],[[positionSpeaker,opinionSpeaker]])
          #print(filteredPermutations)
          #exit()

          #Note: This line should not be set inside the loop on permutations.
          #Reason: It would (wrongly) model the situation that the next speaker is
          #selected depending on the group opinion.
          nextSpeakerPosition = nextSpeakerPoistionAtRandom(positionSpeaker, n)
          for permutation in filteredPermutations:

                #currentTreeTemp will remain 'a' if the permutation was a
                #consensus node.
                currentTreeTemp='a'

```

```
                #Note: in the next line one could probably use
                #the driver as well->
                #filterAll(allPermutations,[filterNotAllAreOne],[[opinionSpeaker]])
                #instead of using a direct call.
                if(filterNotAllAreOne(permutation,[opinionSpeaker])):

                   #Note: this line contains the position nextSpeakerPosition.
                   currentTreeTemp =  generateTreeStructure1_random(0,n,K,permutation[nextSpeakerPosition],distanceToRoot+1,nextSpeakerPosition,allPermutations, distanceToComplete-1)
                #currentTreeTemp=='a' if this was an agreement node with
                #the opinionSpeaker opinion.
                treeStructure.append([permutation,currentTreeTemp,None])

      return treeStructure


#This function returns a random position(uniformly selected) that
#indicates the next speaker. Naturally, a speaker that just spoke can not
#speak again.
def nextSpeakerPoistionAtRandom(positionSpeakerJustSpoke, n):

     nextPositionOfSpeaker = positionSpeakerJustSpoke
     while nextPositionOfSpeaker == positionSpeakerJustSpoke:

          nextPositionOfSpeaker = np.random.randint(0,n)

     return nextPositionOfSpeaker

#######
def isLeaf(treeStructure):

   return (treeStructure[1]=='a' or treeStructure[1]=='d')

#Generates all the opinions at a distance.
#Distance is not meant to be larger than K*n
def getOpinionsAtDistance(treeStructure,distance):
    opinions=[]
    if(distance == 0):
        return [treeStructure[0]]

    else:

       if(isLeaf(treeStructure)): return []
       for tree in treeStructure[1]:

           opinionsTemp= getOpinionsAtDistance(tree,distance-1)
           opinions= opinions + opinionsTemp

       return opinions

###Check for consensus numbers##################
###For every simulation, the following check will take place:
###The agreement-nodes per level that was generated as part of the tree structure,
###needs to be equal to the theoretical one that is known from equations.
###Background, prevent us from having bugs in the generation of the tree structure.
###maxDistance is meant to be K*n at most.
```

```
def checkAgreement(treeStructure,O,n,opinion,maxDistance):
  existAgreement = True
  for distance in range(1,maxDistance+1):
      for o in range(O):
         if o==opinion:
             existAgreement= (existAgreement and checkAgreementAtDistance(treeStructure,O,n,o, True,distance))
         else:
             existAgreement= (existAgreement and checkAgreementAtDistance(treeStructure,O,n,o, False,distance))


         if(not existAgreement):
             print("no agreement between the theorethical consensus node-count and the empirical consensus node-count " )
             print("opinion: "+ str(o))
             print("distance: "+ str(distance))
             exit()
  return  existAgreement

########This function check the same property as the previous function, but
#at a particular distance from the root.
#minimum possible distance is 1.
def checkAgreementAtDistance(treeStructure,O,n,opinion, wasInitial,distance):

  #gets all the empirical opinions at a certain distance from the root
  opinionsAtdistance= getOpinionsAtDistance(treeStructure,distance)

  #calculates the theoretical amount of consensus-nodes at a certain distance
  #from the root, and specifying if the node
  #has the same opinion that the initial speaker had.
  theoreticalConsensusAtDistance= calculateNumberOfConsensusAtDistance(O,n,distance,wasInitial)

  #filter from all the empirical opinions at a distance those
  #that has a given opinion.
  empiricalConsensusAtDistance= filterAll(opinionsAtdistance,[filterAllAreOne],[[opinion]])

  #check that the theoretical and empirical amount of consensus at
  #the given distance are the same.
  return  (theoreticalConsensusAtDistance == len(empiricalConsensusAtDistance))

###This function return the total probability of all agreement-nodes.
###it uses the function evalProbabilities, which does the same(for a given opinion)
def sumProbabilitites(treeStructure,O):
    probSumTotal=0
    for o in range(0, O):
       probSumTotal += evalProbabilities(treeStructure, o)

    return probSumTotal

###this functions return a list with the probabilites of consensus of each opinion
def probabilityNumeratorPerSpeaker(treeStructure,O):
    probPerSpeaker=[]
    for o in range(0, O):
       probPerSpeaker.append( evalProbabilities(treeStructure, o))
```

```

    return probPerSpeaker

###Returns the total probability(should be 1.0)..this is just a way to check
#that the probability assigements are correct.
def checkSumProbabilities(treeStructure,O):

    totalProbability=0
    probabilititesDenominator= sumProbabilitites(treeStructure,O)
    probabilitiesNumerator= probabilityNumeratorPerSpeaker(treeStructure,O)

    for p_i in probabilitiesNumerator:

        totalProbability+=(p_i/probabilititesDenominator)

    return totalProbability

#this function calculates the number of consensus-nodes at a distance i.
#The boolean isFirstSpeaker controls both cases: the consensus was the
#opinion of the first speaker or otherwise.
def calculateNumberOfConsensusAtDistance(O,n,i,isFirstSpeaker):
     H= pow(O,n-2)
     Q= pow(O,n-1)-1
     if isFirstSpeaker:
       numerator= pow(-1,i)*((H*pow(Q,i)*pow(-1,i))+(Q*(-Q+H-1)))


     else:
       numerator=H*pow(-1,i)*(pow(-1,i)*pow(Q,i) + Q)

     return numerator/(Q*(Q+1))
############Assign probabilities to a tree structure#######
############The probability assigment add up to the unit for each branching
############probabilitites on different branchings are independent.
def assignProbabilitiesToTreeStructure(ts,O,n):
    probabilityAssigment=None
    if(isLeaf(ts)):
         #Note: this "doing nothing" means that the assgiment of
         #probabilitites is made in the moment of visiting a node with offspring.
         return

    else:
        probabilityAssigment=generateDirichletProbabilities(pow(O,n-1))
        indexBranch=0
        for child in ts[1]:
            child[2]=probabilityAssigment[indexBranch]
            indexBranch+=1

        for child in ts[1]:
            assignProbabilitiesToTreeStructure(child,O,n)
###Generate probabilities for a uniform Dirichlet distribution(given size of array)
def generateDirichletProbabilities(branchingSize):
   alpha= tuple(1 for _ in range(branchingSize))
   sample=np.random.dirichlet(alpha)

   return sample

```

```
########################Expectation###########
def theoreticalExpectation(n,O,K):

   p=1.0/pow(O,n-1)
   l=K*n
   if(l%2==1):
      expectation=(p+pow(p,l+1))/(p+1)
   else:
      expectation=(p-pow(p,l+1))/(p+1)

   return expectation

###Generate a full permutation(given number of players + number of opinions)###
###This is just the driver
###Range of Opinions will be taken from zero to O-1
def fullPermutation(n,O):
   #Initialization
   fullPermutationList=[]
   #Call to the efactual recursive function
   fullPermutation1(O,[],n,fullPermutationList)
   return fullPermutationList

#O:opinions. Range of Opinions will be taken from zero to O-1
#tempPermutation: will keep the permutation that is under_construction.
#lengthPermutation: keeps the length of the permutation that is still to be built.
#fullPermutationList: keep all the permutation that are built.
def fullPermutation1(O,tempPermutation,lengthPermutation,fullPermutationList):

   if(lengthPermutation==0):
     #print(tempPermutation)
     # the copy function ensure that there are not problems with the
     #references during the backtracking.
     fullPermutationList.append(tempPermutation.copy())

   else:

      for i in range(O):
          tempPermutation.append(i)
          fullPermutation1(O,tempPermutation,lengthPermutation-1,
    fullPermutationList)
          tempPermutation.pop()

   return

###Generate a filtered permutation(from a full permutation + a list of filters)##
def filterAll(permutations,filters,arguments):
   filteredPermutations=[]

   for permutation in permutations:
        fTest=True
        for filterIndex in range(len(filters)):
             fTest = fTest and filters[filterIndex](permutation,arguments[
    filterIndex])

        if(fTest):
            filteredPermutations.append(permutation)
   #print(filteredPermutations)
```

```
   return filteredPermutations

###########Filter1(given a permutation and a particular opinion)#########
###########returns true if not all the speakers have the given opinion###
#Note: the == operators requires types are the same.
#The previous requires to keep an eye on the type of the elements when
#permutations are generated.
def filterNotAllAreOne(permutation,params):
   opinion=params[0]
   for op in permutation:
        if not op==opinion:
            return True

   return False

###########Filter that is the negation of the previous one###
def filterAllAreOne(permutation,params):
   opinion=params[0]
   for op in permutation:
        if not op==opinion:
            return False

   return True

###Filter2(given a permutation, a particular opinion, and a position)
###returns true if the particular permutation has the opinion at the given position
def filterAtPosition(permutation,params):
   position=params[0]
   opinion=params[1]

   return permutation[position]==opinion



#n,O,K
#Note: In the text of the disertation, profiles are
#described with a [O,K,n] structure(and not n,O,K as here).
profiles=[[2,2,1],[2,3,1],[2,4,1],[2,5,1],[2,6,1],[2,7,1],
          [2,8,1],[2,9,1],[2,10,1],[2,11,1],
          [2,2,2],[2,3,2],[2,4,2],[2,5,2],[2,6,2],[2,7,2],
          [2,8,2],[2,9,2],[2,10,2],[2,11,2],
          [2,2,3],[2,3,3],[2,4,3],[2,5,3],[2,6,3],[2,7,3],[2,8,3],[2,9,3],
          [3,2,1],[3,3,1],[3,4,1],[3,5,1],[3,6,1],[3,7,1],[3,8,1],
          [3,2,2],[3,3,2],
          [3,2,3],
          [3,2,4]]
numberOfExperiments=20000
####Experiment1
def experiment1(profiles,numberOfExperiments):
  for profile_index in range(0,len(profiles))  :

    sumAllExperiments=0

    n=profiles[profile_index][0]
    O=profiles[profile_index][1]
    K=profiles[profile_index][2]
    for experiment_i in range(0,numberOfExperiments):
```

```

       st_i =  generateTreeStructure(O,n,K,0,generateTreeStructure1)

       #Note: we work with 0 as the opinion of the first speaker and with 1
       #as the opinion of any other speaker. The use of the particular value 1
       #does not affect the generality of this results because any other
       #opinion will appear on a tree with the same regularity that 1 does.
       S_0=evalProbabilities(st_i, 0)
       S_1=evalProbabilities(st_i, 1)
       sumAllExperiments += (S_0-S_1)


    print([O,K,n],sumAllExperiments/numberOfExperiments, theoreticalExpectation(n,O,K))

#experiment1(profiles,numberOfExperiments)
####Experiment2################################################################
def experiment2(profiles,numberOfExperiments):
  for profile_index in range(0,len(profiles))  :

    s_0StrictlyLargerThanS_1=0

    n=profiles[profile_index][0]
    O=profiles[profile_index][1]
    K=profiles[profile_index][2]
    for experiment_i in range(0,numberOfExperiments):

       st_i =  generateTreeStructure(O,n,K,0,generateTreeStructure1)

       S_0=evalProbabilities(st_i, 0)
       S_1=evalProbabilities(st_i, 1)

       if(S_0>S_1):
          s_0StrictlyLargerThanS_1+=1


    print([O,K,n],s_0StrictlyLargerThanS_1,numberOfExperiments-s_0StrictlyLargerThanS_1,s_0StrictlyLargerThanS_1/numberOfExperiments)

#experiment2(profiles,numberOfExperiments)
####Experiment3################################################################
def experiment3(profiles,numberOfExperiments):
  for profile_index in range(0,len(profiles))  :

    s_0StrictlyLargerThanS_1=0

    n=profiles[profile_index][0]
    O=profiles[profile_index][1]
    K=profiles[profile_index][2]
    for experiment_i in range(0,numberOfExperiments):

       st_i =  generateTreeStructure(O,n,K,0,generateTreeStructure1_random)

       S_0=evalProbabilities(st_i, 0)
       S_1=evalProbabilities(st_i, 1)

       if(S_0>S_1):
```

```
          s_0StrictlyLargerThanS_1+=1


    print([O,K,n],s_0StrictlyLargerThanS_1,numberOfExperiments-s_0StrictlyLargerThanS_1,s_0StrictlyLargerThanS_1/numberOfExperiments)

#experiment3(profiles,numberOfExperiments)
####Experiment4###################################################################
#n,O,K
#This experiment requires that for each profile,
#the number of allowed opinions is a power of two.
"""
profiles=[[2,4,1],[2,8,1],[2,4,2],[2,8,2],
          [2,4,3],[2,8,3],
          [3,4,1],[3,8,1],
          [3,2,4]]
"""
numberOfExperiments=20000

def experiment4(profiles,numberOfExperiments):

  for profile_index in range(0,len(profiles)):

      #This variable keeps the scores per profile.
      #and it is printed after the cyle of experiments
      #corresponding to each profile.
      record_deliberation_profile = [0] * profiles[profile_index][1]



      n=profiles[profile_index][0]
      #This expresses that opinions will be debated in duos.
      O=2
      K=profiles[profile_index][2]
      for experiment_i in range(0,numberOfExperiments):

          #In a scenario with A allowed opinions, the variable
          #current_profile_index_status will keep (initially) indexes
          #from 0 to (A-1)
          #Next, the intention is to reduce the length of
          #current_profile_index_status to
          #its half on each iteration of the following loop.
          #At the end of each experiment, the variable
          #current_profile_index_status will
          #contain the index of the opinion that prevailed.
          #Note: because of the position of the indexes in
          #current_profile_index_status,
          #the opinion with index zero will enjoy of more debates in
          #the first position of the presentation.
          current_profile_index_status = list(range(0,profiles[profile_index][1]))

          debate_index=0
          #This condition expresses the intention:
          #at the end, we should have a single winner opinion.
          while ((len(current_profile_index_status) != 1)) :

              st_i =  generateTreeStructure(O,n,K,0,generateTreeStructure1)
```

```
              S_0=evalProbabilities(st_i, 0)
              S_1=evalProbabilities(st_i, 1)

              if(S_0>S_1):
                   #If the first position won,
                   #the other one is removed.
                   current_profile_index_status.pop(debate_index+1)

              if(S_1>S_0):
                   #If the second position won,
                   #the first one is removed.
                   current_profile_index_status.pop(debate_index)

              if(S_1==S_0):
                   #If none won, one of them is selected(randomly)
                   #and removed.
                   index_to_remove= np.random.randint(0,2)
                   current_profile_index_status.pop(debate_index+index_to_remove)

              debate_index = debate_index+1

              #This condition allows to reset the variable debate_index
              #after each loop of opinion reduction. This is important,
              #otherwise debate_index will be larger than the number of
              #remining opinions.
              if(debate_index+1 > (len(current_profile_index_status)-1)):
                 debate_index =0

          #Once the "while" loop ends, the variable
          #current_profile_index_status will contain a
          #single element(it is in the 0-position).
          #A "victory" will be recorded in the position
          #of that element.
          record_deliberation_profile[current_profile_index_status[0]] +=1

      print([profiles[profile_index][1],K,n],record_deliberation_profile)


#experiment4(profiles,numberOfExperiments)
#################################################################################
#This functions returns a list of probability values.
#Each value is meant to be used as a probability bound
#for the deliberation to end in the first round with a consensus
#equal to the opinion of the first speaker.
#For generating the probability values, the strategy is
#to move around the expectation of that value(in both directions).
#The returned list is ordered.
def getProbabilityBounds(O,n,boundProfileSize):

    #This is the expectation of the probability in a branch(due to the uniform Dirichlet distribution).
    theoretical_exp = 1.0/pow(O,n-1)

    boundValueProfile=[]

    #We return as many values as the dopple
    #as the value of boundProfileSize
```

```
    #Reason: one value to the left and one value
    #to the right.
    for value_profile_size in range(0,boundProfileSize):

            boundValueProfile.insert(value_profile_size,theoretical_exp+(
    theoretical_exp/pow(2,value_profile_size+1)))
            boundValueProfile.insert(value_profile_size,theoretical_exp-(
    theoretical_exp/pow(2,value_profile_size+1)))

    return boundValueProfile

#This function produces a dirichlet probability value(uniform)
#for each branch, with the additional constraint that the
#probability of the first speaker should be bounded above by
#the value contained in the bound parameter.
def getDirichletWithBound(O,n,bound):

    #This generates a dirichlet probability value(uniform) for each branch.
    probabilityAssigment = generateDirichletProbabilities(pow(O,n-1))

    #Without loss of generality, we assume that the first value of the vector
    #will be the value of the first speaker.
    #This condition and the consequent actions enforce the bound.
    #Basically, the remaining probability beyond the bound, is shared
    #among the other edges.
    if(probabilityAssigment[0] > bound):

         #This difference will be equally divided among all the other edges.
         diff_prob =  probabilityAssigment[0]-bound
         share_of_diff_prob = diff_prob/(pow(O,n-1)-1)

         #The probability of the first speaker branch will be set to the
         #max value defined by the bound.
         probabilityAssigment[0] = bound

         loop_index=1
         while(loop_index < len(probabilityAssigment)):
               probabilityAssigment[loop_index] =  probabilityAssigment[loop_index]
     + share_of_diff_prob
               loop_index = loop_index+1

    #If the bound condition does not hold, the initial list of probabilities
    #will be returned.
    return probabilityAssigment

#This function takes two arguments:
#a tree and a new probability assigment
#for the first-level edges of the tree.
#It returns the same tree with the new probabilitites
#in place.
def reAssignProbabilititesFirstLevel(ts,newProbabilityAssigment):

    #This is a loop index variable.
    while_index_reassign_probabilitites = 0

    #There will be a cycle for each subtree.
    while(while_index_reassign_probabilitites < len(ts[1])):
```

```

         #The original values of the first level edges
         #will be set to the values in newProbabilityAssigment
         #Important: The probability value of the edge that goes
         # to the vertex (0,0,0,...,0) is in the first position of
         #newProbabilityAssigment. This only happens because of the
         #order in which permutations are generated in fullPermutation1.
         #If the agrement vertex was (1,1,1,...,1) the probability
         #will not be at the first position of the
         #newProbabilityAssigment list.
         ts[1][while_index_reassign_probabilitites][2] = newProbabilityAssigment[while_index_reassign_probabilitites]
         while_index_reassign_probabilitites = while_index_reassign_probabilitites + 1

    return ts

def experiment5(profiles,numberOfExperiments):

  for profile_index in range(0,len(profiles)):

     n=profiles[profile_index][0]
     O=profiles[profile_index][1]
     K=profiles[profile_index][2]

     #This line produces a list with four elements
     #because for getProbabilityBounds(O,n,boundProfileSize)
     #the given value of boundProfileSize here is 2.
     probBounds = getProbabilityBounds(O,n,2)

     #This means that we not only iterate by profile,
     #but by probability bound too.
     for probValBound in probBounds:

        #This variable will keep track of the
        #results of the experiment.
        s_0StrictlyLargerThanS_1 = 0

        #This ensures that we perform a number
        #of experiments equal to numberOfExperiments.
        for experiment_i in range(0,numberOfExperiments):

           #This line generates the usual opinion tree.
           st_i =  generateTreeStructure(O,n,K,0,generateTreeStructure1)

           #This line gets the new probability values for
           #the first level of edges. The new values ensure
           #that the probability of consensus is bounded
           #(at the first round).
           dirichlet_with_bound = getDirichletWithBound(O,n,probValBound)

           #Next, we set the new probability values
           #in place.
           st_i = reAssignProbabilititesFirstLevel(st_i,dirichlet_with_bound)

           S_0=evalProbabilities(st_i, 0)
           S_1=evalProbabilities(st_i, 1)
```

```

            if(S_0>S_1):
                s_0StrictlyLargerThanS_1+=1


        print([O,K,n,probValBound,1.0/pow(O,n-1)],s_0StrictlyLargerThanS_1,
    numberOfExperiments-s_0StrictlyLargerThanS_1,s_0StrictlyLargerThanS_1/
    numberOfExperiments)

#experiment5(profiles,numberOfExperiments)
#################################################################################
#Note: The following functions are testing functions.
#They can be used to test the correctness of the generated structures:
#getOpinionsAtDistance(treeStructure,distance)
#checkAgreement(treeStructure,O,n,opinion,maxDistance)
#checkAgreementAtDistance(treeStructure,O,n,opinion, wasInitial,distance)
#sumProbabilitites(treeStructure,O)
#probabilityNumeratorPerSpeaker(treeStructure,O)
#checkSumProbabilities(treeStructure,O)
#calculateNumberOfConsensusAtDistance(O,n,i,isFirstSpeaker)
```

Listing D.1: Experiments

Bibliography

Lehrer, K. and C. Wagner (1981). *Rational Consensus in Science and Society: a philosophical and mathematical study*. D. Reidel Pub. Co. ISBN: 9789027713063.

Taylor, J.R. (1997). *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*. Second edition. University Science Books. ISBN: 0935702423.

Bazerman, M.H. (2002). *Judgment in managerial decision making*. 5th ed. New York: Wiley. ISBN: 9780471398875.

Hegselmann, R. and U. Krause (2002). "Opinion Dynamics and Bounded Confidence, Models, Analysis and Simulation". In: *Journal of Artificial Societies and Social Simulation* 5.3. URL: http://jasss.soc.surrey.ac.uk/5/3/2.html.

Mussweiler, T., B. Englich, and F. Strack (2004). "Anchoring Effect". In: *Cognitive illusions: A handbook on fallacies and biases in thinking, judgement and memory*. Ed. by Rüdiger Pohl. Psychology Press. ISBN: 9781841693514.

Nisan, N., ed. (2007). *Algorithmic game theory*. Cambridge University Press. ISBN: 9780521872829.

Hughes, I. and T.P.A. Hase (2010). *Measurements and Their Uncertainties: A Practical Guide to Modern Error Analysis*. Oxford University Press. ISBN: 9780199566327.

Furnham, A. and H.C. Boo (2011). "A literature review of the anchoring effect". In: *The Journal of Socio-Economics* 40.1, pp. 35–42. ISSN: 10535357.

Kadane, J.B. (2011). *Principles of Uncertainty*. First edition. CRC Press. ISBN: 9781439861615.

Ng, K.W., G.L. Tian, and M.L. Tang (2011). *Dirichlet and Related Distributions: Theory, Methods and Applications*. Wiley. ISBN: 9780470688199.

Billingsley, P. (2012). *Probability and Measure, Anniversary Edition*. Wiley. ISBN: 9781118122372.

Estlund, David M., ed. (2012). *The Oxford handbook of political philosophy*. Oxford handbooks. Oxford University Press. ISBN: 9780195376692.

Zhu, D.H. (2013). "Group polarization on corporate boards: Theory and evidence on board decisions about acquisition premiums: Group Polarization on Corporate Boards". In: *Strategic Management Journal* 34.7, pp. 800–822. ISSN: 01432095. (Visited on 06/22/2023).

Leobacher, G. and F. Pillichshammer (2014). *Introduction to quasi-Monte Carlo integration and applications*. Birkhäuser. ISBN: 3319034251.

Tuschke, A., W. Sanders, and E. Hernandez (2014). "Whose experience matters in the boardroom? The effects of experiential and vicarious learning on emerging market entry: Board Ties, Learning, and Emerging Market Entry". In: *Strategic Management Journal* 35.3, pp. 398–418. ISSN: 01432095. (Visited on 06/22/2023).

Malhotra, S., P. Zhu, and T.H. Reus (2015). "Anchoring on the acquisition premium decisions of others: Anchoring in Acquisition Premiums". In: *Strategic Management Journal* 36.12, pp. 1866–1876. ISSN: 01432095. (Visited on 06/22/2023).

Blackburn, S. (2016). *The Oxford dictionary of philosophy*. Third edition. Oxford quick reference. OCLC: ocn942359185. Oxford: Oxford University Press. ISBN: 9780198735304.

Burgman, M.A. (2016). *Trusting judgements: how to get the best out of experts*. Cambridge University Press. ISBN: 9781107112087.

Rubinstein, R.Y. and D.P. Kroese (2016). *Simulation and the Monte Carlo method*. Third edition. John Wiley & Sons, Inc. ISBN: 9781118632161.

Bächtiger, A. et al., eds. (2018). *The Oxford handbook of deliberative democracy*. First edition. Oxford handbooks. OCLC: on1016922115. Oxford University Press. ISBN: 9780198747369.

Durrett, R. (2019). *Probability: Theory and Examples*. Fifth edition. Cambridge University Press. ISBN: 9781108473682.

Frigg, R. and S. Hartmann (2020). "Models in Science". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Spring 2020. Metaphysics Research Lab, Stanford University.

Hartmann, S. and S. Rafiee Rad (Oct. 2020). "Anchoring in Deliberations". en. In: *Erkenntnis* 85.5, pp. 1041–1069. ISSN: 0165-0106, 1572-8420. DOI: 10.1007/s10670-018-0064-y. URL: http://link.springer.com/10.1007/s10670-018-0064-y.

West, D.B. (2020). *Combinatorial Mathematics*. Cambridge University Press. ISBN: 9781107058583.

Cormen, T.H. et al. (2022). *Introduction to algorithms*. Fourth edition. Cambridge, Massachusett: The MIT Press. ISBN: 9780262046305.

List, Ch. (2022). "Social Choice Theory". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta and Uri Nodelman. Winter 2022. Metaphysics Research Lab, Stanford University.

Shabunin, M.I. (original author) (2022). *Asymptotic equality. Encyclopedia of Mathematics*. URL: http://encyclopediaofmath.org/index.php?title=Asymptotic_equality&oldid=32388 (visited on 04/17/2022).

Weisstein, E.W. (2022). *"Condorcet's Jury Theorem". From MathWorld--A Wolfram Web Resource*. en. Text. URL: https://mathworld.wolfram.com/CondorcetsJuryTheorem.html (visited on 11/03/2022).