Deliberation as Evidence Disclosure: A Tale of Two Protocol Types

Extended Abstract

Julian Chingoma
ILLC, University of Amsterdam
The Netherlands
j.z.chingoma@uva.nl

Adrian Haret
ILLC, University of Amsterdam
The Netherlands
a.haret@uva.nl

ABSTRACT

We propose a model inspired by deliberative practice in which agents selectively disclose evidence about alternatives prior to taking a final decision on them. We are interested in whether such a process results in the objectively best alternative getting elected, thereby lending support to the idea that groups can be wise even when their members communicate with each other. We find that, under certain restrictions on the relative amounts of evidence, together with the actions available to the agents, there exist deliberation protocols in each of the two families we look at (i.e., simultaneous and sequential) that offer desirable guarantees. Simulation results further complement this picture, by showing how the distribution of evidence among the agents influences the outcome of the protocols.

KEYWORDS

Deliberation; Social Choice Theory; Epistemic Democracy

ACM Reference Format:


1 INTRODUCTION

Arguments that groups can be wise, i.e., that they can be trusted to find an objectively correct answer, date back to Condorcet [8, 14]. However, work in this tradition typically operates under the assumption that agents are independent, and thereby precluded from communicating with (or otherwise influencing) each other. At the same time, there is much enthusiasm amongst proponents of deliberative democracy around the idea that deliberation, if done right, can lead to better outcomes [4, 15, 17, 20, 22, 24, 34]—which, if correct, would imply that at least some form of communication is beneficial. With the recent proliferation of platforms allowing for widespread sharing of information, it becomes all the more important, for the future of a meaningful democratic process, to determine whether deliberation bolsters (or hinders) the truth-tracking task. The question is whether the exchange of information among members of a group can be structured in a way that is conducive to accurate beliefs and, down the line, to correct decisions.

This is the question we want to address here. We propose a formal model where, in keeping with the epistemic social choice tradition [8, 10, 11, 14, 27], there is an objective ranking over alternatives. This ranking, in our model, is determined by evidence supporting each alternative. Agents rank alternatives based on the evidence they have access to, and ultimately take a collective decision. We take it that evidence may be unevenly distributed among the population, and this leads to varied, possibly incorrect, beliefs. The corrective to an ill-informed opinion, in our model, is communication, here formalized as evidence disclosure: we assume agents are truth-seeking, hence both open to changing their beliefs on the quality of alternatives by absorbing evidence disclosed by others, as well as willing to inform others by selective disclosure of evidence. Formally, this is modelled by a deliberation protocol, which we think of as a process that takes place in rounds and consists of rules about how information is disclosed and processed by the agents. The process stops when no one can, or wishes to, do anything to further change the status quo. We are interested in when such deliberative practices lead to good decisions. Our model’s dynamics bears resemblance to that seen in iterative voting [25, 29, 30] or other deliberation-inspired models [7, 13, 16, 26, 32].

2 THE MODEL

A group of \(n\) agents vote over two alternatives, denoted \(a\) and \(b\). Each alternative is associated with a finite set of evidence items, which determine an objective ranking (more is better). We assume \(a\) is optimal, i.e., there is more overall evidence supporting \(a\) than \(b\). Evidence is distributed among the population, with each agent privy to a subset of the evidence for each alternative. Agents rank alternatives based on their private evidence. Rankings may or may not coincide with the true one, and ties are possible: in particular, it is possible to see the sub-optimal alternative \(b\) elected as the majority winner based on the initial, private evidence vote. In an idealization of the distributions we might expect to see, we assume that evidence for both \(a\) and \(b\) is distributed fully among the agents, and that evidence sets, per alternative per agent, are disjoint.

We study three deliberation protocols, grouped in two families: one simultaneous and two sequential. All protocols work in rounds.
At each round agents decide whether to disclose evidence in support of some alternatives. As soon as evidence is disclosed all agents update their evidence sets and, consequently, their rankings. Intuitively, agents are willing to disclose evidence if they think there are alternatives deserving of more support than the ones currently winning. We distinguish two types of agents, lazy and keen, that support alternatives they think are strictly better than, and weakly better than, any alternative under consideration, respectively. Agents are assumed to disclose only new (i.e., not previously disclosed) evidence; in case they support some alternative and possess new evidence at a round, we say they are dissenting at that round. Protocols terminate when there are no dissenting agents. The plurality winners over the profile obtained at this point are the final winners.

Who discloses, and when? In the simultaneous protocol there is a single disclosure instance per round, with dissenting agents disclosing all private evidence for the alternatives they support (see Figure 1). In both sequential protocols, on which more below, agents take turns based on some fixed ordering to nominate alternatives, with a tally of these nominations used to determine the alternatives currently winning. For the sequential constant protocol the tally is reset at the start of every round; an agent nominates at their turn and may disclose one evidence item for each supported alternative. For the sequential abstention protocol the tally is never reset; on their turn, an agent’s either nominates and discloses an evidence item for their chosen nominee(s), or abstains and does nothing.

The protocols presented here are guaranteed to terminate: ideally, they do so with the optimal alternative as the final winner. When does this happen?

3 MAIN RESULTS

Theoretical guarantees. If the evidence distribution is unbalanced, with small (but very vocal!) sets of supporters for \( b \) swaying the remaining electorate in the wrong direction, the optimal alternative \( a \) can lose out to \( b \) (see Figure 1). We might expect such a situation to be prevented if \( a \) is supported by (much) more evidence than \( b \). Writing \( E(a) \) and \( E(b) \) for the overall amount of evidence for \( a \) and \( b \), respectively, we can make this intuition precise.

**Theorem 1.** If agents are lazy, all deliberation protocols are guaranteed to terminate with \( a \) as the final winner iff \( |E(a)| > n \cdot |E(b)| \).

If agents are keen, this happens iff \( |E(a)| > n \cdot |E(b)| - n \).

Outside of the bounds of Theorem 1, the simultaneous protocol works well when \( a \) begins as the underdog, as its supporters rally to flip the outcome. Sequential protocols are sensitive to the order in which agents speak. So a putative debate organizer privy only to the agents’ rankings (not evidence sets) can influence the result when fixing the order. The next results show when success is achievable.

**Theorem 2.** If there are no neutral agents, then with an agent ordering where agents who put \( b \) at the top speak first, followed by agents who put \( a \) at the top, the sequential constant protocol terminates with \( a \) as final winner.

**Theorem 3.** If all agents are keen, then the sequential abstention protocol terminates with \( a \) as final winner, for any agent ordering.

Simulations. Experiments using randomly generated distributions of evidence across agents suggest that protocols can perform well even when the evidence gap for \( a \) over \( b \) is below the bounds of Theorem 1, if (i) there is enough of an evidence gap for \( a \) over \( b \) (the larger the better), and (ii) the distributions of evidence for \( a \) and \( b \) across agents are relatively similar, with some of the results summarized in Figure 2. What exactly counts as ‘enough’ at (i) deserves scrutiny in future work, while (ii) can be parsed as follows: one common feature of bad outcomes (e.g., Figure 1) is that evidence for \( a \) starts out distributed roughly equally, with all agents getting an equal share of the total evidence for \( a \), whereas evidence for \( b \) is heavily skewed towards a few agents. These agents dissent and derail the final outcome. The hypothesis, borne out by simulations, is that such unbalanced initial distributions spell trouble.

4 CONCLUSIONS

We have put forward a model of information exchange that approximates deliberation processes and that is worth studying and extending further. We found theoretical guarantees if the evidence gap for the optimal alternative is large enough; if not, the moral is that careful orchestration of the rules of debate is needed: sequential protocols can be successful by regulating the order in which agents speak, and by exploiting assumptions on agent behavior (i.e., if they are keen or lazy). Simulations show that, on average, the optimal alternative has a better chance of winning if the evidence distribution for each alternative is similar across agents. Further investigation is needed here to uncover meaningful parameters.

Our model is not probabilistic, i.e., we do not touch upon the process by which agents initially acquire evidence. Nonetheless, adding on a model of belief formation is an obvious avenue for future work, with the aim of recovering familiar jury theorems [8, 9, 12, 21, 23, 28, 31, 33]. Equally interesting would be to follow the lead of the AI literature on opinion dynamics [1–3, 5, 6, 18, 19] by restricting agent communication to local neighborhoods.
ACKNOWLEDGMENTS

We wish to thank our anonymous AAMAS reviewers for their valuable feedback.

REFERENCES


