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:
Julian Chingoma and Adrian Haret. 2023. Deliberation as Evidence Disclosure: A Tale of Two Protocol Types: Extended Abstract. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 3 pages.

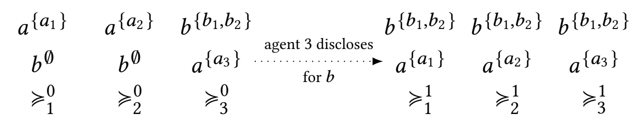

Figure 1: **Alternatives a and b are supported by evidence $\{a_1, a_2, a_3\}$ and $\{b_1, b_2\}$, respectively. Based on the evidence they have (depicted as superscripts), agents 1, 2 and 3 start out ranking a and b differently ($\succcurlyeq_i^t$ is i's ranking at round t). Under the simultaneous protocol, at round 0 we have that agent 3 (regardless of type), the only agent unahppy with the plurality opinion at this point, discloses *all* their available evidence for b. Doing so makes 1 and 2 change their minds and update their rankings. At round 1 alternative b is the plurality winner, with no agent disagreeing any further. The protocol ends here. The same outcome occurs under the two sequential protocols.**

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.

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom.

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 neither.

The protocols presented here are guaranteed to terminate: ideally, they do so with the optimal alternative a 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 $|E(a)| \geqslant n \cdot |E(b)|$. If agents are keen, this happens iff $|E(a)| \geqslant n \cdot |E(b)| - n$.

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

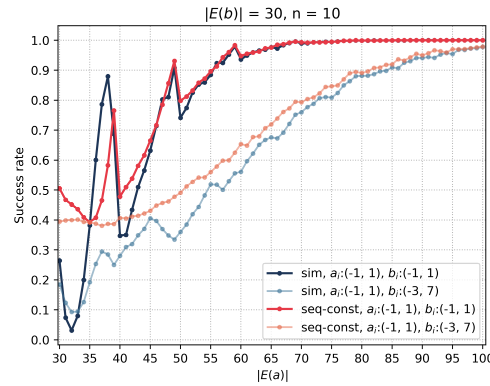

Figure 2: We fix the number n of agents and overall amount $|E(b)|$ of evidence for b, and vary the amount $|E(a)|$ of evidence for a. Success rate measures the percentage of times a ends up as the winner over 5000 instances. Note that success rate grows larger as the evidence gap in favor of a increases. Second, we vary the *spread* in the distributions of evidence, summarized by the amount by which evidence can deviate from the average, e.g., for $n = 10$ agents and $|E(a)| = 30$, an equal distribution of evidence sees every agent get $30/10 = 3$ items of evidence for a; $a_i : (-1, 1)$ means that each agent actually has in between 2 and 4. Intuitively, a larger spread means that some agents can end up holding large amounts of evidence at the expense of other agents. Note that larger variance for b relative to a leads to a lower success rate.

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

[1] Daron Acemoglu, Asuman Ozdaglar, and Ali ParandehGheibi. 2010. Spread of (mis)information in social networks. *Games and Economic Behavior* 70, 2 (2010), 194–227.

[2] Vincenzo Auletta, Ioannis Caragiannis, Diodato Ferraioli, Clemente Galdi, and Giuseppe Persiano. 2015. Minority Becomes Majority in Social Networks. In *Proc. of the International Conference on Web and Internet Economics (WINE 2015)*. Springer, 74–88.

[3] Vincenzo Auletta, Diodato Ferraioli, Valeria Fionda, and Gianluigi Greco. 2019. Maximizing the Spread of an Opinion when Tertium Datur Est. In *Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019)*, Edith Elkind, Manuela Veloso, Noa Agmon, and Matthew E. Taylor (Eds.). IFAAMAS, 1207–1215.

[4] André Bächtiger, John S. Dryzek, Jane Mansbridge, and Mark E. Warren. 2018. *The Oxford Handbook of Deliberative Democracy.* Oxford University Press.

[5] Sushil Bikhchandani, David Hirshleifer, Omer Tamuz, and Ivo Welch. 2021. Information Cascades and Social Learning. (June 2021).

[6] Markus Brill, Edith Elkind, Ulle Endriss, and Umberto Grandi. 2016. Pairwise Diffusion of Preference Rankings in Social Networks. In *Proc. of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI 2016)*, Subbarao Kambhampati (Ed.). IJCAI/AAAI Press, 130–136.

[7] Hun Chung and John Duggan. 2020. A Formal Theory of Democratic Deliberation. *American Political Science Review* 114, 1 (2020), 14–35.

[8] Marie Jean Antoine Nicolas de Caritat, Condorcet. 1785. Essai Sur l'application de l'analyse à La Probabilité Des Décisions Rendues à La Pluralité Des Voix. (1785).

[9] Franz Dietrich and Kai Spiekermann. 2013. Epistemic Democracy With Defensible Premises. *Economics & Philosophy* 29, 1 (2013), 87–120.

[10] Franz Dietrich and Kai Spiekermann. 2020. Jury Theorems. In *The Routledge Handbook of Social Epistemology*, Miranda Fricker, Peter J. Graham, David Henderson, and Nikolaj J.L.L. Pedersen (Eds.). Routledge, 386–396.

[11] Franz Dietrich and Kai Spiekermann. 2021. Social Epistemology. In *The Handbook of Rationality*, Markus Knauff and Wolfgang Spohn (Eds.). MIT Press, 579–590.

[12] Franz Dietrich and Kai Spiekermann. 2022. Deliberation and the Wisdom of Crowds. *Documents de travail du Centre d'Économie de la Sorbonne* (2022).

[13] Huihui Ding and Marcus Pivato. 2021. Deliberation and epistemic democracy. *Journal of Economic Behavior & Organization* 185 (2021), 138–167.

[14] Edith Elkind and Arkadii Slinko. 2016. Rationalizations of Voting Rules. In *Handbook of Computational Social Choice*, Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D. Procaccia (Eds.). Cambridge University Press, 169–196.

[15] Jon Elster. 1998. *Deliberative Democracy.* Cambridge University Press.

[16] Brandon Fain, Ashish Goel, Kamesh Munagala, and Sukolsak Sakshuwong. 2017. Sequential Deliberation for Social Choice. In *Proc. of the International Conference on Web and Internet Economics (WINE 2017)*, Nikhil R. Devanur and Pinyan Lu (Eds.). Springer, 177–190.

[17] James S. Fishkin and Peter Laslett. 2003. *Debating Deliberative Democracy.* Blackwell.

[18] Benjamin Golub and Matthew O. Jackson. 2010. Naïve Learning in Social Networks and the Wisdom of Crowds. *American Economic Journal: Microeconomics* 2, 1 (2010), 112–149.

[19] Benjamin Golub and Evan Sadler. 2016. Learning in Social Networks. In *The Oxford Handbook of the Economics of Networks*, Yann Bramoullé, Andrea Galeotti, and Brian W Rogers (Eds.). Oxford University Press, 503–542.

[20] Robert E. Goodin and Kai Spiekermann. 2018. *An Epistemic Theory of Democracy.* Oxford University Press.

[21] Bernard Grofman, Guillermo Owen, and Scott L. Feld. 1983. Thirteen Theorems in Search of the Truth. *Theory and Decision* 15, 3 (1983), 261–278.

[22] Stephan Hartmann and Soroush Rafiee Rad. 2018. Voting, deliberation and truth. *Synthese* 195, 3 (2018), 1273–1293.

[23] Krishna K. Ladha. 1992. The Condorcet Jury Theorem, Free Speech, and Correlated Votes. *American Journal of Political Science* (1992), 617–634.

[24] Hélène Landemore. 2013. *Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many.* Princeton University Press.

[25] Omer Lev and Jeffrey S. Rosenschein. 2012. Convergence of iterative voting. In *Proc. of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012)*, Wiebe van der Hoek, Lin Padgham, Vincent Conitzer, and Michael Winikoff (Eds.). IFAAMAS, 611–618.

[26] Christian List. 2007. Group Deliberation and the Transformation of Judgments: An Impossibility Result. (March 2007).

[27] Christian List. 2018. Democratic Deliberation and Social Choice. In *The Oxford Handbook of Deliberative Democracy*, André Bächtiger, John S. Dryzek, Jane Mansbridge, and Mark E. Warren (Eds.). Oxford University Press.

[28] Christian List and Robert E. Goodin. 2001. Epistemic Democracy: Generalizing the Condorcet Jury Theorem. *Journal of Political Philosophy* 9, 3 (2001), 277–306.

[29] Reshef Meir. 2017. Iterative Voting. In *Trends in Computational Social Choice*, Ulle Endriss (Ed.). AI Access, Chapter 4, 69–86.

[30] Reshef Meir, Maria Polukarov, Jeffrey Rosenschein, and Nicholas Jennings. 2010. Convergence to Equilibria in Plurality Voting. *Proc. of the AAAI Conference on Artificial Intelligence (AAAI 2010)* 24, 1, 823–828.

[31] Guillermo Owen, Bernard Grofman, and Scott L. Feld. 1989. Proving a distribution-free generalization of the Condorcet Jury Theorem. *Mathematical Social Sciences* 17, 1 (1989), 1–16.

[32] Juan Perote-Peña and Ashley Piggins. 2015. A model of deliberative and aggregative democracy. *Economics & Philosophy* 31, 1 (2015), 93–121.

[33] Marcus Pivato. 2017. Epistemic democracy with correlated voters. *Journal of Mathematical Economics* 72 (2017), 51–69.

[34] Marcus Pivato. 2019. Realizing Epistemic Democracy. In *The Future of Economic Design*, Jean-François Laslier, Hervé Moulin, M Remzi Sanver, and William S Zwicker (Eds.). Springer, Cham, 103–112.