

# Who has the last word? Understanding How to Sample Online Discussions

Gioia Boschi  
Department of Mathematics, King's  
College London  
London, UK  
gioia.boschi@kcl.ac.uk

Anthony P. Young  
Department of Informatics, King's  
College London  
London, UK  
peter.young@kcl.ac.uk

Sagar Joglekar  
Department of Informatics, King's  
College London  
London, UK  
sagar.joglekar@kcl.ac.uk

Chiara Cammarota  
Department of Physics, Sapienza  
Università di Roma  
Rome, Italy  
Department of Mathematics, King's  
College London  
London, UK  
chiara.cammarota@kcl.ac.uk

Nishanth Sastry  
Department of Computer Science,  
University of Surrey  
Guildford, UK  
Department of Informatics, King's  
College London  
London, UK  
n.sastry@surrey.ac.uk

## ABSTRACT

In online debates, as in offline ones, individual utterances or arguments support or attack each other, leading to some subset of arguments (potentially from different sides of the debate) being considered more relevant than others. However, online conversations are much larger in scale than offline ones, with often hundreds of thousands of users weighing in, collaboratively forming large trees of comments by starting from an original post and replying to each other. In large discussions, readers are often forced to sample a subset of the arguments being put forth. Since such sampling is rarely done in a principled manner, users may not read all the relevant arguments to get a full picture of the debate from a sample. This paper is interested in answering the question of how users should sample online conversations to selectively favour the currently justified or accepted positions in the debate. We apply techniques from argumentation theory and complex networks to build a model that predicts the probabilities of the normatively justified arguments given their location in idealised online discussions of comments and replies which we represent as trees. Our model shows that the proportion of replies that are supportive, the distribution of the number of replies that comments receive, and the locations of comments that do not receive replies (i.e., the “leaves” of the reply tree) all determine the probability that a comment is a justified argument given its location. We show that when the distribution of the number of replies is homogeneous along the tree length, for acrimonious discussions (with more attacking comments than supportive ones), the distribution of justified arguments depends on the parity of the tree level which is the distance from the root expressed as number of edges. In supportive discussions, which have

more supportive comments than attacks, the probability of having justified comments increases as one moves away from the root. For discussion trees which have a non-homogeneous in-degree distribution, for supportive discussions we observe the same behaviour as before, while for acrimonious discussions we cannot observe the same parity-based distribution. This is verified with data obtained from the online debating platform Kialo. By predicting the locations of the justified arguments in reply trees, we can therefore suggest which arguments readers should sample, to grasp the currently accepted opinions in such discussions. Our models have important implications for the design of future online debating platforms.

## CCS CONCEPTS

• Information systems → World Wide Web; • Human-centered computing → Collaborative and social computing theory, concepts and paradigms.

## KEYWORDS

argumentation theory, online discussions, probabilistic analysis, graph sampling, Kialo

## ACM Reference Format:

Gioia Boschi, Anthony P. Young, Sagar Joglekar, Chiara Cammarota, and Nishanth Sastry. 2022. Who has the last word? Understanding How to Sample Online Discussions. In *Companion Proceedings of the Web Conference 2022 (WWW '22 Companion)*, April 25–29, 2022, Virtual Event, Lyon, France. ACM, New York, NY, USA, 1 page. <https://doi.org/10.1145/3487553.3524188>

---

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

*WWW '22 Companion, April 25–29, 2022, Virtual Event, Lyon, France*

© 2022 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-9130-6/22/04.

<https://doi.org/10.1145/3487553.3524188>