

Are Two Heads Better Than One? An Exploration of Ambiguity in Crowd-Collected Language Decisions from the Phrase Detectives Game.

Jon Chamberlain

School of Computer Science and Electronic Engineering
University of Essex, Colchester, Essex UK, jchamb@essex.ac.uk

ABSTRACT

The online game-with-a-purpose Phrase Detectives (<https://www.phrasedetectives.com>) [1] has been collecting decisions about anaphoric coreference in human language for over 10 years (4 million judgements from 40,000 players). The game was originally designed to collect multiple valid solutions for a single task, which complicated aggregation but created a very rich (and noisy) dataset [2]. Analysis of the ambiguous player decisions highlights the need for understanding and resolving disagreement that is inherent in language interpretation. This talk will present some of the interesting cases of ambiguity found by the players of Phrase Detectives (a dataset that will be made available to the research community later this year [3]) and discuss the statistical methods we have been working on to harness crowds that disagree with each other [4, 5].

ACM Reference format:

Jon Chamberlain. 2019. Are Two Heads Better Than One? An Exploration of Ambiguity in Crowd-Collected Language Decisions from the Phrase Detectives Game.. In *Proceedings of WWW '19: The Web Conference (WWW '19), May 13, 2019, San Francisco, USA*. ACM, New York, NY, USA, 1 page. <https://doi.org/10.1145/3308560.3317311>

1 Speaker biography

Dr Jon Chamberlain is a lecturer and applied research scientist based at the University of Essex, England. He has been the lead developer behind the Phrase Detectives game-with-a-purpose for over 10 years and is co-investigator on the ERC Advanced Grant Disagreements and Language Interpretation (DALI) project that builds on early work to understand ambiguity in human language.

2 ACKNOWLEDGMENTS

The speaker is funded by an ERC Advanced Grant – [ERC-2015-AdvG](#).

This paper is published under the Creative Commons Attribution 4.0 International (CC-BY 4.0) license. Authors reserve their rights to disseminate the work on their personal and corporate Web sites with the appropriate attribution.

WWW '19, May 13, 2019, San Francisco, USA

© 2019 IW3C2 (International World Wide Web Conference Committee), published under Creative Commons CC-BY 4.0 License.

ACM ISBN 978-1-4503-6675-5/19/05.

<https://doi.org/10.1145/3308560.3317311>

REFERENCES

- [1] Poesio, M., Chamberlain, J., Kruschwitz, U., Robaldo, L., and Ducceschi, L. (2013). Phrase Detectives: Utilizing collective intelligence for internet-scale language resource creation. *ACM Transactions on Interactive Intelligent Systems*, 3(1):1-44.
- [2] Chamberlain, J., Poesio, M., and Kruschwitz, U. (2016). Phrase Detectives corpus 1.0: Crowdsourced anaphoric coreference. In *Proceedings of LREC, Portoroz, Slovenia*.
- [3] Poesio, M., Chamberlain, J., Paun, S., Kruschwitz, U., and Yu, J. (2019) forthcoming). "A Crowdsourced Corpus of Multiple Judgments and Disagreement on Anaphoric Interpretation." In *Proceedings of The 17th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2019)*, Minneapolis. Association for Computational Linguistics.
- [4] Paun, S., Carpenter, B., Chamberlain, J., Hovy, D., Kruschwitz, U., and Poesio, M. (2018). "Comparing bayesian models of annotation." *Transactions of the Association for Computational Linguistics*.
- [5] Paun, S., Chamberlain, J., Kruschwitz, U., Yu, J., and Poesio, M. (2018). "A probabilistic annotation model for crowdsourcing coreference." In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 1926-1937, Brussels, Belgium. Association for Computational Linguistics.