Unjust Equivalence: Are Irony and Sarcasm Truly the Same in NLP?

Florijan Sandalj, Bojan Puvača, Ivan Unković

University of Zagreb, Faculty of Electrical Engineering and Computing
Unska 3, 10000 Zagreb, Croatia
{florijan.sandalj, bojan.puvaca, ivan.unkovic}@fer.hr

Abstract

1. Introduction

2. Irony and sarcasm in NLP

The relationship between irony and sarcasm is unfortunately a heavily contested subject in NLP. This problem is the easiest to see when looking at different sarcasm and irony datasets, where we can find cases when they are treated as completely seperate concepts (Nikhil, 2020), when sarcasm is treated as a subset of irony (Van Hee et al., 2018) or even vice versa (Oprea and Magdy, 2020) Searching for the definite answer in the realm of linguistics is a futile effort as well, however we have found two distinctions between the two that have merit in the context of NLP.

2.1. Sarcasm - irony's meaner cousin

The online Merriam-Webster dictionary defines sarcasm as "a sharp and often satirical or ironic utterance designed to cut or give pain" (Merriam-Webster, 2024). This definition seems to be in line with the general consensus that sarcasm is a form of irony that is more aggressive and mean-spirited. The iSarcasm dataset (Oprea and Magdy, 2020) is a good example of this categorization, as the "sarcasm" label is, in fact, a subset of the unfortunately named "sarcastic" label, which actually indicates any kind of ironic speech.

2.2. Sarcasm - the figure of speech

3. Experimental setup

3.1. Sarcasm detection dataset

(Oprea and Magdy, 2020)

4. Results

5. Discussion

6. Conclusion

Acknowledgements

References

Merriam-Webster. 2024. Merriam-webster dictionary.

John Nikhil. 2020. Tweets with sarcasm and irony.

Silviu Oprea and Walid Magdy. 2020. iSarcasm: A dataset of intended sarcasm. In Dan Jurafssky, Joyce Chai, Natalie Schluter, and Joel Tetreault, editors, *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 1279–1289, Online, July. Association for Computational Linguistics.

Cynthia Van Hee, Els Lefever, and Véronique Hoste. 2018. SemEval-2018 task 3: Irony detection in English tweets. In Marianna Apidianaki, Saif M. Mohammad, Jonathan May, Ekaterina Shutova, Steven Bethard, and Marine Carpuat, editors, *Proceedings of the 12th International Workshop on Semantic Evaluation*, pages 39–50, New Orleans, Louisiana, June. Association for Computational Linguistics.