

Reading Assignment 4

COMP 550, Fall 2024

Due date: November 28, 2024 9:00pm

Read the following paper and write a critical summary of it:

Rudinger et al., 2018. Gender Bias in Coreference Resolution. *NAACL*.
<https://arxiv.org/pdf/1804.09301>

Your summary should include a description of the paper's contents, its strengths and limitations, and any points that you did not understand. Some aspects of this paper may not be covered in the lectures, I expect that there will be concepts that you do not understand! Do not worry, and document this in your write-up.

Also, discuss the following issues in your writeup:

1. How does this paper propose to measure gender bias in coreference resolution systems? Explain how an "unbiased" system should perform, versus the results obtained in the experiments.
2. The paper discusses a potential case of bias amplification involving the occupation *manager* on page 4. Explain the proposed mechanism in which dataset bias could be amplified at the system level, then how system bias could then lead to further amplification in society downstream.
3. Why does the paper say that their work has low negative predictive value?

Your summary should be **at most 1 page**. It will be graded on the basis of its coverage, linguistic quality, and clarity of argumentation.

Resources

- Jason Eisner's guide to reading papers:
<https://www.cs.jhu.edu/~jason/advice/how-to-read-a-paper.html>
- A closely related paper that you may want to compare against:
<https://arxiv.org/pdf/1804.06876>