Supplementary material

Chenhao Tan[*1], Hao Peng[†2], and Noah A. Smith[‡2]

[1]Department of Computer Science, University of Colorado Boulder
[2]Paul G. Allen School of Computer Science & Engineering, University of Washington

1   Quote Extraction.

In this work, we differentiate quotes from quotations. We refer to any texts in news articles that are enclosed in quotation marks as *quotations*, and *quotes* are a subset of quotations that can be matched to a turn at a debate. A quotation *matches* a turn in a debate if either more than 80% of 3-grams in the quotation are in the turn, or the quotation fuzzy matches a substring of the turn.[1]

In order to extract pairs of quotes and quoted sentences for each debate, we first identify relevant news articles, then extract all quotes in these relevant news articles that have at least three words, and finally match quotes to quoted sentences in the debate.

In general, it is a non-trivial task to decide whether a news article is related to a debate. To avoid false positives, we define *relevant news articles* as articles with at least a 5-word quotation that can be matched to a turn from the corresponding debate. We also remove verbatim transcripts of presidential debates from our news article collection.

After we identify relevant news articles for a debate, we include all quotes that have at least three words in those articles to exclude short common

---

[*]chenhao@chenhaot.com

[†]hapeng@cs.washington.edu

[‡]nasmith@cs.washington.edu

[1]We use a Levenshtein distance based metric (`https://github.com/seatgeek/fuzzywuzzy`). We require "ratio" as defined in the package to be larger than 90, roughly suggesting more than 90% of character overlap.
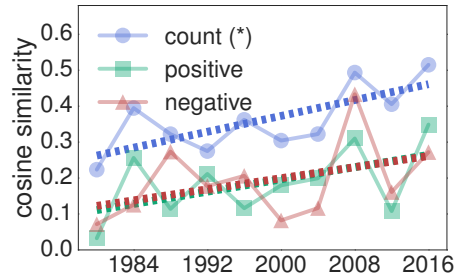
Figure 1: Similarity between the New York Times and Washington Post. * in the legend indicates that the correlation between y-values and time is statistically significant with $p < 0.05$.

quotes that may not come from the debate (e.g., "good" can be in quotation marks) and avoid confusion in identifying the quoted sentence.

Finally, we match each quote to quoted sentences in the corresponding debate. As a quote sometimes consists of parts from multiple sentences, we consider parts that are separated by "..." in quotes as basic units and match each basic unit to a sentence in the debate.

2 Similarity between the New York Times and Washington Post

We show the comparison between the two largest media outlets in quoting general debates as a case study. In fact, the New York Times and Washington Post are often the nearest neighbor with each other. The similarity is increasing over time, which is consistent with the overall trend in similarity at the local level. The three lines are not always moving together: comparing 2000 to 1996, although the New York Times and Washington Post used positive context in more similar ways, but it is not the case for counts and negative context. We do not see as prominent a spike in 2016, in fact, NYT and WP demonstrated the same level of similarity in 2008 when Obama first ran against McCain.

3 Human Performance in Figure 4(a)

In our Mechanical Turk experiments, although each turker labeled 5 random pairs from each threshold in {2, 4, 6, 8, 10}, the human performance in

Figure 4(a) is estimated based on subsets of pairs in the 200 random pairs for $t = 10$.