

Readings

I found out that the title of last week’s weekly reading, ‘Conversations gone alright’ (Bao et al. 2021), is an homage to an early paper entitled ‘Conversations gon awry: Detecting early signs of conversation failure’ (Zhang et al. 2018).

Zhang et al. (2018) construct computational models to predict (based on an initial post/comment in a conversational exchange) whether a conversation with descend into anti-social behaviour. They note that politeness has been shown in the literature to be important to how conversations develop. Their research questions have significant similarity to one of our questions. We are asking about the textual and schematic markers of advice-seeking posts w.r.t. to whether the conversation proceeds civilly.

Some findings we can incorporate into our hypotheses and research questions are that ‘hedged remarks’ in the initial conversation prompt are associated with sustained civility, whereas forceful questions are associated with descending from civility. Zhang et al. (2018) not that ‘direct language addressing the other interlocutor’ is associated with lesser civility, which further motivates our hypothesis that use of the 2nd person pronoun as subject is associated with conversations developing badly.

Processing and Analysis

I finished manually annotating the 100 contrived sentences and 100 example comments from r/AITA. We discussed the fact that the 100 example comments were lacking examples of thankfulness. To rectify this we decided to sample 50 more second level comments for manual annotation, half of which we ensured contained the string ‘thank’. I wrote the script for extracting all top-level and second-level comments for each submission, from which Pranav did the sampling of 50 more comments.

We were originally planning to sample just 10,000 submissions for RQ1, but we have now changed our methodology to the following:

We first find all submission names associated with submissions that have neither been deleted by the user (`selftext` = “[*deleted*]”) nor removed by a moderator (`selftext` = “[*removed*]”), which brings us from 1,702,055 down to 468,649 submissions. A rather large number of submissions have been either deleted or removed, which is perhaps to be expected given the forum is frequently used for intimate self-disclosure. of these 468,649 submissions we uniformly sample 50,000 for further processing. Next we gather all of the comments associated with each of the 50,000 submissions, extracting them from `AmItheAsshole_comments.txt` based on the `parent_id` field matching the `name` of the submission. This results in a collection of 5,262,806 unique comments.

To answer RQ1 we need examples of advice-seeking submissions with a negative comment (namely where the comment contains either YTA or ESH). Each comment is directly descended from one other node in the discussion tree. For

this dataset we are interested in comments descended directly from the first submission, i.e. the root of the tree. We say the root node (the submission) has depth 0, $d = 0$, a comment directly on the submission has depth 1, $d = 1$, and so on. Comments with depth 1 are named TLCs (Top Level Comments) by Bao et al. (2021). We call comments with depth 1 2LCs (2nd Level Comments). Secondly, we need to see that the OP has responded to the negative comment. Therefore a sample in this dataset has three texts, the original submission c_0 authored by a_1 ($d = 0$), a comment c_1 on the original submission containing either of the substrings `YTA'' orESH` authored by $a_2 \neq a_1$ ($d = 1$), and a comment c_2 authored by a_1 ($d = 2$) on the comment c_1 authored by a_2 .

After processing the 50,000 submissions and 5,262,806 associated comments we are left with 12,649 triplets following the above structure.

References

Bao, Jiajun, Junjie Wu, Yiming Zhang, Eshwar Chandrasekharan, and David Jurgens. “Conversations gone alright: Quantifying and predicting prosocial outcomes in online conversations.” In Proceedings of the Web Conference 2021, pp. 1134-1145. 2021.