Neural-based RST Parsing And Analysis In Persuasive Discourse

Jinfen Li
School of Information Studies,
Syracuse University
`jli284@syr.edu`

Lu Xiao
School of Information Studies,
Syracuse University
`lxiao04@syr.edu`

Abstract

Most of the existing studies of language use in social media content have focused on the surface-level linguistic features (e.g., function words and punctuation marks) and the semantic level aspects (e.g., the topics, sentiment, and emotions) of the comments. The writer's strategies of constructing and connecting text segments have not been widely explored even though this knowledge is expected to shed light on how people reason in online environments. Contributing to this analysis direction for social media studies, we build an openly accessible neural RST parsing system that analyzes discourse relations in an online comment. Our experiments demonstrate that this system achieves comparable performance among all the neural RST parsing systems. To demonstrate the use of this tool in social media analysis, we apply it to identify the discourse relations in persuasive and non-persuasive comments and examine the relationships among the binary discourse tree depth, discourse relations, and the perceived persuasiveness of online comments. Our work demonstrates the potential of analyzing discourse structures of online comments with our system and the implications of these structures for understanding online communications.

1   Introduction

Researchers are increasingly interested in analyzing the language used in social media content. Previous studies examine surface-level linguistic features such as the use of function words or punctuation marks in online communications (Tan et al., 2016), and/or semantic level aspects such as understanding the sentiment (Neri et al., 2012) and emotions (Li and Xiao, 2020a) as well as identifying topics (Nzali et al., 2017). Recently, studies also conduct discourse-level language analysis such as exploring cohesion in the social media content (Latifah and Triyono, 2020) and using discourse level structure for detecting fake news (Karimi and Tang, 2019). The pragmatic use of the language in the content is also limitedly explored, such as the analysis of imperative statements in Wikipedia discussions (Xiao and Nickerson, 2019). Arguing for the importance of considering context information in language analysis, Benamara et al. (2018) call for more NLP research that "develop tools to automatically determine the structure of discourse, including discourse relations, argumentation, and threads in conversations such as those found in Twitter and other social media" (Benamara et al., 2018, p. 676).

As a response to this call, we have developed a discourse parsing system that automatically constructs the discourse parsing tree for the input text. Discourse parsing is a task of parsing a text segment into a set of discourse units and logically connect the merged units using a set of discourse relations, which exhibit reasoning behaviors in online environments (Öcal et al., 2021). Rhetorical Structure Theory (RST) (Mann and Thompson, 1988) is one of the most influential theories in discourse parsing and is used in our system. To parse RST, we need to build up a tree, where the nodes bear information such as span, nuclearity, and relation. The leaf node in the RST tree represents an Elementary Discourse Unit (EDU), a sequence of which could be merged into a longer span. The span place in the non-terminal nodes attaches with the nuclearity and relation labels. Nuclearity type typically includes NN, NS and SN, is the semantic role in discourse, where the nucleus span (N) is more essential than the satellite span (S).

The neural-based system of (Kobayashi et al., 2020) utilizes granularities including paragraphs, sentences and EDUs levels. The paragraphs and sentences granularities forcedly split a span at the paragraph/sentence boundary instead of other positions, resulting in the wrong subspans. There are around 7.6% EDUs with inconsistent sentence or paragraph index in one EDU. In addition, they

parse RST in a top-down manner, which results in a larger search space. For instance, there are $n - 1$ possible split points for a span composed of n EDUs. They also add a second-layer BLSTM upon all the EDUs to view bidirectional context. Building on their work, we conduct the following tasks to achieve a higher performance in RST parsing:

- We emphasize the role of position by introducing the sinusoid positional embedding and pay attention to the first and last two tokens in an EDU.

- We explore whether an additional layer of BLSTM, extracting correlations among EDUs in the document levels is helpful in such a bottom-up system.

- We apply different components in the relation detection submodel based on different nuclearity types.

To demonstrate the use of our discourse parsing system for analyzing discourse relations in social media content, we apply it in online persuasion research in two online contexts, namely, Wikipedia AfD and Reddit CMV discussions. We investigate the differences between persuasive and non-persuasive comments regarding their discourse structure. Our statistical analyses show that a comment's perceived persuasiveness is correlated to discourse relations and the depth of its binary discourse tree. Not only showing the potential of our discourse parsing system for social media content analysis, but these findings also suggest that studying the discourse structure of social media content achieves a better understanding of complex online interactions such as persuasion.

2 Related Works

Previous studies in RST parsing have two dominant parsing approaches including transition-based and neural-based approaches. The transition-based systems that rely on hand-crafted features have achieved great success (Wang et al., 2017; Li and Xiao, 2020b); while neural-based approaches do not require feature engineering, which does not take into account the particular nature of a specific data. Networks such as attention-based hierarchical neural network (Lin et al., 2019), and pointer networks (Lin et al., 2019) are developed to parse RST at sentence or text level. However, neither do

neural network approaches consider word position and nor they design components for different nuclearity types, as initial and end words and nuclearity type are essential in RST parsing (Li et al., 2014; Wang et al., 2017).

The Discourse parsing tools provide a new perspective in discourse interpretation in social media (Benamara et al., 2018). Son et al. (2018) conduct causal explanations in physical and mental health building on discourse parsing; Zakharov et al. (2020) develop a novel discourse annotation schema that reflects a hierarchy of discursive strategies for the analysis of contentious and polarizing discussions. Discourse parsing also is leveraged in sentiment analysis (Bhatia et al., 2015), identify authorship (Ferracane et al., 2017) and detect deception (Rubin and Lukoianova, 2015). In this paper, we build up a new comparable discourse parsing tool for the contribution to social media content analysis in the discourse aspect.

3 RST Parsing System

In this section, we present our RST parsing system. We propose a neural RST parsing system in the bottom-up manner utilizing EDU granularity. We emphasize the role of word positions, investigate the role of an additional layer of BLSTM, and design different relation submodels based on different nuclearity types.

3.1 Representation for an EDU

The RST parsing tree is built over the smallest units: EDUs, whose representation is generated by our encoder. For each EDU, we concatenate the vectors from ELMo (Peters et al., 2018), GloVe vector (Pennington et al., 2014) and the sinusoid positional encoding (PE) (Vaswani et al., 2017) as the embedding of i-th token, i.e., emb_i. We then pass emb_i to the BLSTM layer, and gain the bidirectional hidden states h_i.

$$emb_i = [ELMo(t_i); GloVe(t_i); PE(t_i)] \quad (1)$$

$$h_i = BLSTM(h_{i-1}, emb_i) \quad (2)$$

Previous studies have found that the first and last two words are significant when detecting discourse relation between two spans (Li et al., 2014). We come up with two new context vectors to focus only on the first and last two words in an EDU. As Equation 3 is shown, c_{first} focus on the first hidden

state and c_{last} focus on the last two hidden states in an EDU. We use $cGate$ to denote how much attention should be paid to these two contexts. We choose the max product of $cGate$ and h_i as the representation of an EDU. The overview of the representation of an EDU is shown in Figure 1.

$$c_{first} = h_0, c_{last} = [h_{n-2}; h_{n-1}] \\ cGate = U_{first}c_{first} + U_{last}c_{last} \tag{3}$$

$$h'_i = h_i \odot cGate \\ e_t = \max(h'_0, ..., h'_{n-1}) \tag{4}$$

In Kobayashi et al. (2020)'s study, they concatenate the first and last hidden states as the context vector c'. They propose a weight matrix W_c for the hidden states and another weight matrix U_c for c'. To prove the effectiveness of the partial attention in an EDU, we use their $cGate'$ in the baseline model and $cGate$ in ours.

$$c' = [h_{n-1}; h_0] \\ cGate' = \sigma(W_c h_i + U_c c') \tag{5}$$
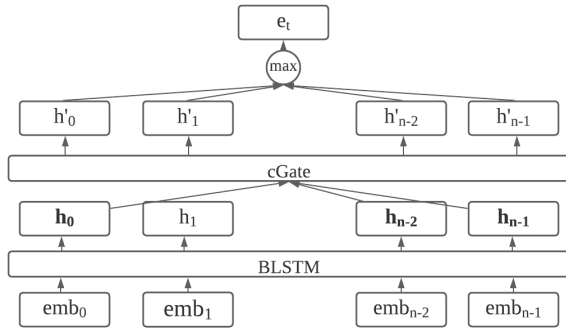

Figure 1: Representation of an EDU

3.2 Representation for a Span

Kobayashi et al. (2020) added an additional BLSTM layer upon all the EDUs for global information benefits. While the additional BLSTM layer provides the global context in their top-down parsing system, it is unknown that whether such context is helpful in our bottom-up parsing system. We directly use $\overrightarrow{e_t}$ and $\overleftarrow{e_t}$ as the forward and backward representation of the t-th EDU.

The boundary features are effective in the three granularities RST parsing system (Kobayashi et al., 2020). We also utilize four kinds of boundary information: whether the span is the start of a sentence, the end of a sentence, the start of a paragraph, or

the end of a paragraph. We concatenate the difference of the forward and backward representation between the left and right spans, and the boundary feature B which is shown in Equation 6.

$$f_t = \overrightarrow{e_t}, b_t = \overleftarrow{e_t} \\ s_{m:n} = [f_n - f_{m-1}; b_{m-1} - b_n; B_{m:n}] \tag{6}$$

3.3 Our Discourse Parsing System

The transition-based system is widely adopted in the bottom-up parsing systems, which proposes a series of actions to build a right-heavy binary tree. The queue stores the unprocessed EDUs and the stack stores the processed ones. Initially, all the EDUs are in the queue. An EDU will be shifted from queue to stack and the top two elements in the stack will be reduced into a subtree. We propose three submodels to detect action, nuclearity, and relation separately.

We first use the single feedforward layer MLP^*_{left} and MLP^*_{right} activated by ReLU function upon the left (from m-th to k-th EDUs) and right (from $k+1$ to n-th EDUs) span vectors, where * denotes that MLP is trained separately in the submodels.

$$s'_{m:k} = MLP^*_{left}(s_{m:k}) \\ s'_{k+1:n} = MLP^*_{right}(s_{k+1:n}) \tag{7}$$

3.3.1 Action Submodel

Yu et al. (2018) choose the top three subtrees in a stack and the first EDU in the queue as the clue for prediction. Instead of predicting the bare action, they predict the information-rich actions (e.g., Reduce-SN-Elaboration) to identify span, nuclearity, and relation in the same stage. In our study, we only predict the bare actions (Shift and Reduce) based on the top two elements in the stack.

For a left span and a right span, we calculate the action score using Equation 8, where W is a weight matrix.

$$action(m, n, k) = W^a_{left}s'_{m:k} + W^a_{right}s'_{k+1:n} \tag{8}$$

3.3.2 Nuclearity Submodel

We decide to use another submodel to predict nuclearity, considering the inferior performance of predicting the joint actions composed of bare action and nuclearity (e.g., Reduce-SN). If the predicted action is reduced, we will detect nuclearity and relation for the reduced subtree. Similar to

the action submodel, we come up with two weight matrices W_{left}^n and W_{right}^n corresponding to $s'_{m:k}$ and $s'_{k+1:n}$. We predict one nuclearity type from NN, NS, and SN.

$$nuc(m, n, k) = W_{left}^n s'_{m:k} + W_{right}^n s'_{k+1:n} \quad (9)$$

3.3.3 Relation Submodel

The nuclearity type is essential in detecting relation in RST parsing (Wang et al., 2017). We take advantage of the nuclearity type predicted by the nuclearity submodel and design two kinds of weight matrices: W_{nuc} and W_{sat}, where W_{nuc} aims to extract information from nucleus spans and W_{sat} learns to exploit information from satellite spans.

Due to the imbalanced relation labels (please check the details in the Appendix A), we separately train two relation submodels for the high frequency and low-frequency labels. Specifically, we train the first relation submodel on the labels including "Attribution", "Enablement", "Elaboration", "Joint", "Same-Unit", and the second submodel on the rest of the labels.

$$rel(m, n, k) =$$
$$\begin{cases} W_{nuc} s'_{m:k} + W_{nuc} s'_{k+1:n}, nuc = NN \\ W_{nuc} s'_{m:k} + W_{sat} s'_{k+1:n}, nuc = NS \\ W_{sat} s'_{m:k} + W_{nuc} s'_{k+1:n}, nuc = SN \end{cases}$$
$$(10)$$

To test the effectiveness of our designed weight matrices based on different nuclearity types in the relation submodel, we apply the following equation in all the compared models except our final model.

$$rel'(m, n, k) = W_{left}^r s'_{m:k} + W_{right}^r s'_{k+1:n}$$
$$(11)$$

Due to the different converge speeds of different submodels, we calculate the loss score using the equation as follows.

$$Loss = \alpha_1 L_{action} + \alpha_2 L_{nuc} + \alpha_3 L_{rel1} + \alpha_4 L_{rel2}$$
$$(12)$$

4 Experiments

4.1 Setup

We build up our bottom-up neural parsing model by using the dataset RST-DT (Mann and Thompson, 1988) (347 and 38 documents in training and test dataset). We split 40 documents from the training dataset to be our development dataset, by using the split method from Heilman and Sagae (2015). 18 coarse relations are explored in our study.

For the hyper-parameters, we use the small size model of ELMo. The initial learning rate is 0.001 and it is decayed with the ratio of 0.8 in each epoch. The gradient clipping threshold is 5.0 and the weight decay is 0.005. The hidden size, positional embedding size, boundary feature size are 200, 30, 10. α_1, α_2, α_3 and α_4 are set to 0.5, 0.3, 0.3 and 0.25. We train the model for 10 epochs and choose the one with the highest span score on the development dataset, and evaluate it on the test dataset.

4.2 Model Comparison

We compare our models with the current state-of-the-art neural bottom-up parsing system of YZF18 and one of the top-down system of D2E.

YZF18 Yu et al. (2018)'s study parses RST based on the transition system, with implicit syntax features generated by an extra neural syntax parser.

D2E We compare Kobayashi et al. (2020)'s model in EDU granularity. D2E is a partial model from Kobayashi et al. (2020) that utilizes EDU level of granularity, while their full model uses three levels of granularity in a document. Their full model includes the preprocess such as forcedly split a span at the paragraph boundary. It would be fairer to compare our model with their full model using the same preprocess, which is not used in our parsing tool. Therefore, we only compare the one without such preprocessing, i.e., D2E.

Model	PE	Equation 3/4	second BLSM	Equation 10/11
baseline	✗	4	✓	11
position	✓	4	✓	11
part	✓	3	✓	11
BLSTM_1	✓	3	✗	11
final	✓	3	✗	10

Table 1: Configurations of compared models. ✓ means inclusion, ✗ means exclusion and BLSTM_1 means single layer of BLSTM.

Our models We conduct an ablation study by changing one condition based on the previous modified model. The configurations are shown in Table 1.

4.3 Result

We present the results of the compared models in Table 2. Our baseline model outperforms *D2E* system on span detection, which implicates our

bottom-up system is more capable of building up the tree structure. Comparing *position* model with the baseline, the positional embedding is helpful in span, nuclearity, and relation detection. The partial attention of EDUs adopted in *part* model further improves the F1 score of nuclearity and relation, by around 1.5. The removal of additional BLSTM upon all the EDUs in a document does not further improve the nuclearity detection but improves the span and relation detection slightly. Our final model, with nuclearity type awareness in the relation submodel, attains the comparable performance. We outperform *YZF18* on the F1 score of relation by 0.4 and achieve 87.9 and 77.3 on span and nuclearity detection.

Model	Span	Nuclearity	Reletion
D2E	86.1	73.1	58.9
YZF18	85.5	73.1	60.2
baseline	86.5	70.3	55.1
position	86.7	71.6	56.9
part	86.7	73.0	58.4
BLSTM_1	86.8	72.3	59.0
final	**87.9**	**73.6**	**60.6**
Human	88.3	77.3	65.4

Table 2: Micro averaged F1 scores based on RST-Parseval.

5   Discourse Relations in Social Media Content: Case Studies

We apply our RST parsing system to identify the discourse structure of a comment in two online discussions, namely, Wikipedia AfD and Reddit CMV discussion datasets. These two datasets have been used in online persuasion studies (Xiao and Xiao, 2020; Xiao and Khazaei, 2019). Both datasets have comments labeled as persuasive or non-persuasive. We choose these datasets because we are interested in whether and how the discourse structure of a comment correlates with its perceived persuasiveness. Persuasion is commonly considered as "human communication that is designed to influence others by modifying their beliefs, values, or attitudes" (Tan et al., 2016; Xiao and Khazaei, 2019). In Iyer et al. (2017) 's study, the researchers found that the syntactic structure of a sentence plays an important role in indicating different persuasion tactics. However, the discourse structure of a com-

ment remains unstudied. The first step in our analysis is to segment the comment to EDUs for the study. We segment a comment by sentence, i.e., the EDU in our case study is a sentence. Comments that only have one sentence are not included in the analysis.

- **Wikipedia AfD dataset** The dataset of Wikipedia "Article for Deletion" is collected by (Mayfield and Black, 2019), which includes the discussion that whether the focal article should be deleted in the Wikipedia community. It allows users to provide their opinion and rationale that whether to keep the article in the discussion that lasts about a couple of weeks. 16,957 comments are labeled persuasive and 23,813 non-persuasive. Of these data, 1,698 persuasive comments and 5,832 non-persuasive ones are removed in our analysis as they only contain one sentence.

- **Reddit CMV dataset** The dataset of Reddit "Change My View" includes users' posts and their comments from February 2015 to January 2016. The one we adopted is collected by Xiao and Khazaei (2019), which only focuses on the direct replies to the post, to address the concern that comments in a thread receiving a delta due to the conversational context, not the persuasive aspects. 1,690 comments are labeled persuasive and 1,690 non-persuasive. 44 persuasive and 93 non-persuasive comments are removed from the analysis because they only have just one sentence.

We identify the discourse structure of a comment in these two datasets using our RST parsing system and then investigate the relationships among a comment's binary discourse tree depth, discourse relations, and perceived persuasiveness. We refer to binary tree depth as shallow and deep with the value of 0 and 1 respectively in the following sections. In these discourse structure analyses, we divide the 18 raw discourse relations into three groups according to the definition of Presentational, Subject_matter, and Multinuclear relations proposed by Mann and Thompson (1988). Of these three types of RST relations, we consider Presentational relations to have the strongest persuasion intent as the intended effect of these relations to increase some inclination in the reader such as the reader's degree of agreement or positive view regarding the

Presentational	Background, Enablement, Summary, Explanation
Subject_matter	Attribution, Cause, Condition, Elaboration, Evaluation, Manner-Means,Topic-Comment
Multinuclear	Comparison, Contrast, Joint, Same-Unit, Temporal, Textual-Organization, Topic-Change

Table 3: RST Relations Lists

nucleus, i.e., the main passage of the statement. The purpose of Subject_matter relations is to help the reader recognize the relation or the connection, i.e., to understand the writer's statement. Multinuclear relations are those that the adjacent text segments are equally important in conveying the meaning of the statement. The use of Multinuclear relations reflects a certain rhetoric strategy that the writer applies such as to convey a point by contrasting two things. The full list of groups is presented in Table 3.
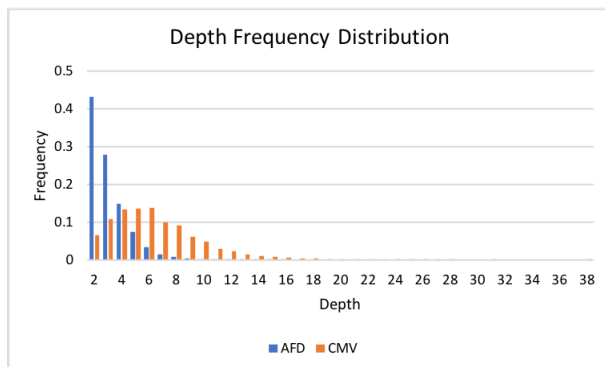

Figure 2: Depth Distribution In Wikipedia Data.

5.1 Qualitative Analysis

To gain a qualitative understanding of how the tool performs on social media data, the two co-authors independently annotate the discourse tree structure for ten randomly selected examples from AfD and CMV datasets and compare the results with those from the tool. An illustration of an example is presented in Figure 3. The two researchers agree on four examples both in terms of their hierarchical structure and their discourse relations. Among these four examples, the tool generates the same annotations for three. These results suggest that the tool has a comparable performance with respect to the annotations that are readily agreed by human experts.

5.2 Is Binary Discourse Tree Depth Relevant To Discourse Relations?

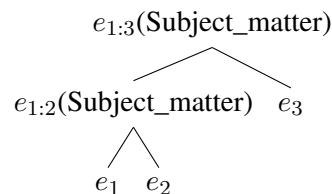In this subsection, we aim to shed light on the relationship between depth and discourse relations.

Figure 3: An illustration of the RST of an example from CMV.
e_1: I do not think you understand why the Christ myth theory is treated as a conspiracy theory;
e_2: All there needs to be in order for Christ to exist is a first-century carpenter named josh who had a few followers and was executed.
e_3: If you find this story plausible you do not believe the Christ myth theory.

It is noted that a longer text does not necessarily mean a deeper binary discourse tree structure. For instance, the average text length in a document with 18 sentences is longer than that with 17 sentences in AfD. As shown in Figure 2, most binary discourse trees have a depth of less than 7. Based on this observation, we define binary tree with depth less than 7 as shallow and deep otherwise. For each discourse tree, we calculate the sum of the frequencies of each corresponding main RST type relation, i.e., Presentational, Subject_matter, or Multinuclear. We conduct the Chi-square test to compare whether the shallow and deep structures are correlated with the type of discourse relations. We present the observed values with expected values in parentheses for each main RST type relations respectively in AfD and CMV in Table 4. Taking 0.05 as the significant level, the results suggest that the depth and the main type of RST relation are significantly correlated in both discussion contexts. The bold elements(bigger actual value than expected value) reveals that discourse tree structure tends to be deep in Presentational and Multinuclear relations, and tend to be shallow in Subject_matter relation. The patterns of correlations between discourse relation and depth hold the same in AfD and CMV, which suggests such correlations are context-independent.

It is expected that Multinuclear relations tend to

	Wikipedia				Reddit			
	Present	Subject	Multi	p value	Present	Subject	Multi	p value
shallow	2603(2790.9)	**57500(56911.5)**	10318(10718.5)	3.23E-58	491(596.5)	**6482(6060.1)**	1640(1956.4)	9.19E-30
deep	**595(407.0)**	7713(8301.4)	**1964(1563.4)**		**1820(1714.5)**	16995(17416.9)	**5939(5622.6)**	

Table 4: Chi-square Test Between Depth And Discourse Relation.

have a deeper discourse tree structure, as our model builds up a binary RST tree instead of a flat multi branches tree. A parent node with four children bearing a multinuclear relation originally will be transformed into a binary tree with a depth of three. The opposite correlation relations that the binary discourse tree depth exhibits with Presentational and Subject_matter relations implies that the binary tree depth may reflect the persuasion intent to some extent. Specifically, as mentioned earlier, Presentational relations reflect the writers' intent to increase belief or acceptance of the nucleus in the reader (Mann and Thompson, 1988), whereas Subject_matter relations mainly have the purpose to make oneself clear(Mann and Thompson, 1988). Therefore, their different correlation behavior with the binary tree depth may be leveraged in tasks like automatically identifying the persuasion intent of social media comments.

5.3 Is Persuasiveness Related To Discourse Relations?

In this subsection, we investigate the potential relationships between a comment's perceived persuasiveness and the type of RST relations it contains. The results of our Chi-square tests show that such a correlation relationship does not exist in Wikipedia AfD discussions (p = 0.3290) but exists in Reddit CMV discussions. We present the actual frequency outside and expected value within the parenthesis in Reddit CMV in Table 5. In the table, the larger observed value in bold elements implies that comments labeled as persuasive are more than expected to have Presentational and Multinuclear relations and less than expected to have Subject_matter relations.

The finding that correlation between persuasiveness and discourse relations exists in CMV discussions but not in AfD discussions implies the context-dependency of a comment's perceived persuasiveness. The discourse relations contained in a comment reflect the writer's use of rhetorical strategies in communication. CMV discussions are about challenging people's viewpoints and trying to

persuade them successfully. Persuasion acts are explicitly expected from individual comments and the goal is to win the argument. The rhetorical strategies applied in constructing the comment are therefore expected to be influential on the comment's perceived persuasiveness. As described earlier, the use of Presentational relations implies the persuasion intent and Multinuclear relations the focus of rhetorical strategies. Subject_matter relations, on the other hand, focus on making the main point understood by the reader. Therefore, it is reasonable to expect that Presentational and Multinuclear relations are more influential on the perceived persuasiveness than Subject_matter relations in CMV discussions.

For Wikipedia AfD discussions, the goal is to identify the best solution on how to handle a Wikipedia article so that the quality of the online encyclopedia is ensured. There is a direct consequence at the end of the discussion that affects the community. Additionally, Wikipedia has a strong online community identity and participants are expected to share this collective goal. AfD discussion focuses on the comment's actual suggestion and its impact is highly considered in evaluating the comment. Hence, the use of rhetorical strategies in a comment is not that influential in the decision-making.

Prior studies have shown that contextual factors affect interaction dynamics in social media communications such as participant entry-order and degree of back-and-forth exchange (Tan et al., 2016; Xiao, 2018). Our finding here presents empirical evidence of the contextual influence on the perceived effects of social media comments' discourse structures.

5.4 Is Discourse Tree Deeper In A Persuasive Comment?

In this subsection, we explore the correlation between persuasiveness and discourse relation in discussion comments. We conduct Spearman's correlation and present the p-value and coefficient in Table 6. We observe a significant correlation between

	Present	Subject	Multi	p value
persuasive	**1497(1475.3)**	14772(14987.4)	**5032(4838.3)**	1.97E-07
non-persuasive	814(835.7)	**8705(8489.6)**	2547(2740.7)	

Table 5: Chi-square Test Between Persuasiveness And Discourse Relation In Reddit CMV Discussions.

perceived persuasiveness and depths, but a different pattern in these two discussions. In Wikipedia AfD discussions, a comment's perceived persuasiveness is negatively correlated to the depth of its binary discourse tree, while in Reddit discussions, the deeper the discourse tree, the more persuasive perceived by the readers.

We calculate the average percentages of three types of relations and find that Subject_matter relation is dominant in both Wikipedia AfD and Reddit CMV discussions. Additionally, they tend to appear more in persuasive comments in the Wikipedia discussions and more in non-persuasive comments in the Reddit discussions. As mentioned in Section 5.2, Subject_matter relations are more likely to appear in a shallow tree structure. Therefore, it explains why a Wikipedia AfD comment's perceived persuasiveness is negatively correlated with the depth of its binary discourse tree and it is the opposite for a Reddit CMV comment.

	Wikipedia	Reddit
p value	0.012	0
coefficient	-0.014	0.22
Subject_matter ratio[per]	0.87	0.73
Subject_matter ratio[non]	0.86	0.77

Table 6: Analysis Between Persuasiveness And Depth.Subject_matter ratio[per] and Subject_matter ratio[non] represent average ratio of Subject_matter relation in persuasive and non-persuasive comments.

6 Discussion

Complex social interactions increasingly happen through social media (Dey et al., 2017; Nguyen et al., 2015). Text's discourse structure reflects how the writer constructs arguments and reasons, which is expected to help us better understand the dynamics of these social interactions. Through the analysis of the comment's discourse structure in the two online discussions and the finding that a comment's discourse structure, identified from the tool, correlates with the comment's perceived persuasiveness, we showcase the potential of using this tool in social media content analysis. Essentially, the availability of this tool offers new opportunities for social media research. How does a comment's discourse structure correlate with the amount of attention it receives, in Twitter, Facebook, or Reddit discussions? Do people converge at the level of discourse structure as the online discussion goes on? Do toxic comments or disinformation tend to have particular discourse structures? These are a few example research questions for these new opportunities.

Our tool currently takes one sentence as an EDU. One of our future tasks is to parse a comment into smaller discourse units and build an RST tree upon them. Also, our model is trained on Wall Street Journal (RST-DT) (Mann and Thompson, 1988), the context and the text length of which is less similar to social media communications. The token number in a sentence in Wall Street Journal (around 18 tokens) is much less than that in AfD (around 63 tokens) and in CMV (around 246 tokens). We are exploring other RST annotated datasets as the training data for our future model.

Our case study of the comment's discourse structure in Wikipedia AfD and Reddit CMV discussions contributes to online persuasion research. The findings of our statistical analyses demonstrate that at the discourse structure level valuable information can be obtained that reflects the persuasion dynamics and helps interpret the perceived persuasiveness of online comments. Additionally, the contextual influence on social media communications can be explored at this level of language analysis as well. We also try to apply the machine learning model proposed by Li and Xiao (2020b) to the case study. The model proposed by Li and Xiao (2020b) achieves the micro F1 score of 87.9, 74.9, 62.6 in span, nuclearity, and relation. The findings remain the same after applying both their and our models, indicating that the performance of an RST parser has little influence on the correlations we reveal in these two online environments. We call for more investigations on how discourse relations and perceived persuasiveness interact in different online

communities contexts.

7 Conclusion

Various natural language processing techniques have been applied in social media content analysis such as word frequency measures, syntactic structure analysis, sentiment analysis, emotion recognition, or topic modeling. Recently, researchers call for further consideration of contextual factors in the analysis and suggest the exploration of the discourse structure (Benamara et al., 2018). In response to this call, we build a bottom-up transition-based system for discourse parsing using the Rhetorical Structure Theory (RST) framework. In this system, we consider the position of tokens in an EDU and We pay partial attention to an EDU regarding the whole semantic of an EDU. We also design relation submodel based on nuclearity types. Our final model achieves the highest performance among all the reported neural network-based systems for RST parsing. We find that the low-frequency relations (please check Appendix) such as Topic-change and Topic-Comment do not achieve satisfying performance. In addition, these relations might benefit from global information. Therefore, we plan to deal with the imbalanced data by paying more attention to the low-frequency relations, and try to embed global information in the bottom-up parsing system in the future.

To demonstrate the use of our model to conduct discourse structure analysis in social media content, we explore the relationships among a comment's binary discourse tree depth, the discourse relations it contains, and its perceived persuasiveness in two online discussion contexts: Wikipedia AfD and Reddit CMV discussions. Our analyses show that the depth of a comment's binary discourse tree and its discourse relations are correlated and independent of the discussion context. In Wikipedia AfD discussion, comments that have Subject_matter relations are more likely to be perceived as persuasive, while in Reddit CMV discussions, comments constructed by Presentational and Multinuclear relations are perceived to be more persuasive.

References

Farah Benamara, Diana Inkpen, and Maite Taboada. 2018. Introduction to the special issue on language in social media: exploiting discourse and other contextual information. *Computational Linguistics*, 44(4):663–681.

Parminder Bhatia, Yangfeng Ji, and Jacob Eisenstein. 2015. Better document-level sentiment analysis from rst discourse parsing. *arXiv preprint arXiv:1509.01599*.

Sanorita Dey, Brittany Duff, Karrie Karahalios, and Wai-Tat Fu. 2017. The art and science of persuasion: not all crowdfunding campaign videos are the same. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pages 755–769.

Elisa Ferracane, Su Wang, and Raymond Mooney. 2017. Leveraging discourse information effectively for authorship attribution. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 584–593.

Michael Heilman and Kenji Sagae. 2015. Fast rhetorical structure theory discourse parsing. *arXiv preprint arXiv:1505.02425*.

Rahul R Iyer, Katia P Sycara, and Yuezhang Li. 2017. Detecting type of persuasion: Is there structure in persuasion tactics? In *CMNA@ ICAIL*, pages 54–64.

Hamid Karimi and Jiliang Tang. 2019. Learning hierarchical discourse-level structure for fake news detection. *arXiv preprint arXiv:1903.07389*.

Naoki Kobayashi, Tsutomu Hirao, Hidetaka Kamigaito, Manabu Okumura, and Masaaki Nagata. 2020. Top-down rst parsing utilizing granularity levels in documents. In *AAAI*, pages 8099–8106.

Ani Latifah and Sulis Triyono. 2020. Cohesion and coherence of discourse in the story of" layangan putus" on social media facebook. *Indonesian Journal of EFL and Linguistics*, 5(1):41–56.

Jinfen Li and Lu Xiao. 2020a. Emotions in online debates: Tales from 4forums and convinceme. *Proceedings of the Association for Information Science and Technology*, 57(1):e255.

Jinfen Li and Lu Xiao. 2020b. Tree representations in transition system for rst parsing. In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 6746–6751.

Sujian Li, Liang Wang, Ziqiang Cao, and Wenjie Li. 2014. Text-level discourse dependency parsing. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 25–35.

Xiang Lin, Shafiq Joty, Prathyusha Jwalapuram, and M Saiful Bari. 2019. A unified linear-time framework for sentence-level discourse parsing. *arXiv preprint arXiv:1905.05682*.

William C Mann and Sandra A Thompson. 1988. Rhetorical structure theory: Toward a functional theory of text organization. *Text*, 8(3):243–281.

Elijah Mayfield and Alan W Black. 2019. Analyzing wikipedia deletion debates with a group decision-making forecast model. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):1–26.

Federico Neri, Carlo Aliprandi, Federico Capeci, Montserrat Cuadros, and Tomas By. 2012. Sentiment analysis on social media. In *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pages 919–926. IEEE.

Duyen T Nguyen, Laura A Dabbish, and Sara Kiesler. 2015. The perverse effects of social transparency on online advice taking. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, pages 207–217.

Mike Donald Tapi Nzali, Sandra Bringay, Christian Lavergne, Caroline Mollevi, and Thomas Opitz. 2017. What patients can tell us: topic analysis for social media on breast cancer. *JMIR medical informatics*, 5(3):e7779.

Ayşe Öcal, Lu Xiao, and Jaihyun Park. 2021. Reasoning in social media: insights from reddit "change my view" submissions. *Online Information Review*.

Jeffrey Pennington, Richard Socher, and Christopher D Manning. 2014. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1532–1543.

Matthew E Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. 2018. Deep contextualized word representations. *arXiv preprint arXiv:1802.05365*.

Victoria L Rubin and Tatiana Lukoianova. 2015. Truth and deception at the rhetorical structure level. *Journal of the Association for Information Science and Technology*, 66(5):905–917.

Youngseo Son, Nipun Bayas, and H Andrew Schwartz. 2018. Causal explanation analysis on social media. *arXiv preprint arXiv:1809.01202*.

Chenhao Tan, Vlad Niculae, Cristian Danescu-Niculescu-Mizil, and Lillian Lee. 2016. Winning arguments: Interaction dynamics and persuasion strategies in good-faith online discussions. In *Proceedings of the 25th international conference on world wide web*, pages 613–624.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in neural information processing systems*, pages 5998–6008.

Yizhong Wang, Sujian Li, and Houfeng Wang. 2017. A two-stage parsing method for text-level discourse analysis. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 184–188.

Lu Xiao. 2018. A message's persuasive features in wikipedia's article for deletion discussions. In *Proceedings of the 9th International Conference on Social Media and Society*, pages 345–349.

Lu Xiao and Taraneh Khazaei. 2019. Changing others' beliefs online: Online comments' persuasiveness. In *Proceedings of the 10th International Conference on Social Media and Society*, pages 92–101.

Lu Xiao and Jeffrey Nickerson. 2019. Imperatives in past online discussions: Another helpful source for community newcomers? In *Proceedings of the 52nd Hawaii International Conference on System Sciences*.

Yimin Xiao and Lu Xiao. 2020. Effects of anonymity on comment persuasiveness in wikipedia articles for deletion discussions. In *Proceedings of the Fourth Workshop on Natural Language Processing and Computational Social Science*, pages 104–115.

Nan Yu, Meishan Zhang, and Guohong Fu. 2018. Transition-based neural rst parsing with implicit syntax features. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 559–570.

Stepan Zakharov, Omri Hadar, Tovit Hakak, Dina Grossman, Yifat Ben-David Kolikant, and Oren Tsur. 2020. Discourse parsing of contentious, non-convergent online discussions. *arXiv preprint arXiv:2012.04585*.

A   Example Appendix

Relation	Count	Relation	Count
Attribution	357	Explanation	126
Background	87	Joint	382
Cause	83	Manner-Means	24
Comparison	37	Same-Unit	342
Condition	27	Summary	16
Contrast	176	Temporal	94
Elaboration	949	Textual-Organization	40
Enablement	62	Topic-Change	12
Evaluation	66	Topic-Comment	20

Table 7: Count of each relation in the test dataset