

Argumentation Mining in Parliamentary Discourse

Nona Naderi^(✉) and Graeme Hirst

University of Toronto, Toronto, Canada
{nona,gh}@cs.toronto.edu

Abstract. We examine whether using frame choices in forum statements can help us identify framing strategies in parliamentary discourse. In this analysis, we show how features based on embedding representations can improve the discovery of various frames in argumentative political speech. Given the complex nature of the parliamentary discourse, the initial results that are presented here are promising. We further present a manually annotated corpus for frame recognition in parliamentary discourse.

1 Introduction

In parliamentary discourse, politicians expound their beliefs and ideas through argumentation, and to persuade the audience, they highlight some aspect of an issue, which is commonly known as framing [5]. Consider the following example¹:

Example 1. *There was no need to change the definition of marriage in order for gays and lesbians to establish meaningful, long term relationships that are recognized in law.*

The speaker is framing his/her argument to promote the idea that the same-sex marriage is not necessary because the long-term relationships are already recognized in law.

While a deep understanding and analysis of beliefs and ideas remains a challenge, a relatively simpler task of argument tagging based on pre-existing frames has recently been proposed [1, 6].

In this paper, we introduce our supervised model trained on user-generated Web content for classifying parliamentary discourse by its use of various frames. We use vector representations of words and sentences to capture their semantic information, and compute semantic similarity metrics across argumentative discourse-pairs. We further present our corpus of argumentative parliamentary discourse. These argumentative speeches are annotated with known frames.

¹ An example from our manually annotated corpus of parliamentary statements.

2 Related Work

We briefly summarize prior work on argument analysis of user postings. Cabrio and Villata [3] used a textual entailment approach to find pro and con arguments in a set of debates selected from Debatepedia². Boltužić and Šnajder [1] proposed a categorization task of tagging user postings with a pre-existing set of frames. Their supervised classification model made use of entailment and semantic similarity features. To generalize their earlier work for various topics, Boltužić and Šnajder [2] presented an unsupervised model to recognize frames by means of textual similarity. In a similar task, Hasan and Ng [6] employed a probabilistic approach for stance and reason classification of user postings. Misra et al. [9] took a supervised approach to classify dialogue postings by “argument facets” using lexical and semantic similarity features. These approaches focused on user-generated content on online forums. In contrast, we explore framing strategies in parliamentary discourse.

3 Corpus and Annotation

For our frame prediction task, we use user-postings manually annotated with known frames (ComArg corpus) as a training set and argumentative parliamentary speeches as a test set. The corpora that we conducted our study on are described in the following sections.

3.1 The ComArg Corpus

ComArg³, developed by Boltužić and Šnajder [1], is a corpus of user statements manually annotated with users’ positions towards a specific topic (pro or con stance), and a set of pre-existing “arguments”. These arguments are, in effect, *frames* in the sense that we introduced above, as each highlights certain aspects of the issue. The authors chose two different sources for collecting their data; the user statements are compiled from ProCon.org, where the statements are associated with a labeled *pro* or *con* stance, and the frames are taken from Idebate.org. The corpus covers two topics of *gay marriage (GM)* and *Under God in Pledge (UGIP)*. Since the latter (regarding the Pledge of Allegiance) is an issue specific to the United States, we focused solely on the GM part of the corpus, which contains 198 statements and 7 pre-existing frames, shown in Table 1⁴. In this corpus, the pairs of statements and frames are annotated as *explicit attack*, *implicit attack*, *no mention*, *explicit support*, and *implicit support*; that is the statements *for* gay marriage can support the *pro* frames, and attack the *con* frames, and vice versa for statements opposing gay marriage. In this work, we only used the statements that explicitly (176 instances) and implicitly (98 instances) *supported* the pre-existing frames.

² <http://idebate.org/debatabase>.

³ <http://takelab.fer.hr/data/comarg/>.

⁴ The third frame is modified to accommodate frames in our current corpus.

Table 1. ComArg pre-defined frames on gay marriage

Frame	Stance	Description
1	Con	Gay couples can declare their union without resort to marriage
2	Pro	Gay couples should be able to take advantage of the fiscal and legal benefits of marriage
3	Con	Gay marriage undermines the institution of marriage
4	Pro	It is discriminatory to refuse gay couples the right to marry
5	Con	Major world religions are against gay marriages
6	Pro	Marriage is about more than procreation; therefore gay couples should not be denied the right to marry due to their biology
7	Con	Marriage should be between a man and a woman

3.2 Argumentative Parliamentary Statements

For our test set, we focused on debates regarding same-sex marriage in the Canadian Parliament. In 2005, Bill C-38, *An act respecting certain aspects of legal capacity for marriage for civil purposes*, to legalize same-sex marriage in Canada, was introduced in the Parliament. Later that year, the bill was passed and the legal definition of marriage was expanded under the then-Liberal government to include conjugal couples of the same sex. After the Conservative Party of Canada gained power, the debate on same-sex marriage was re-opened in the Parliament in 2006; therefore, the issue was debated extensively in the Parliament in two different periods of time (same-sex marriage was debated briefly in 1999).

We selected speeches regarding same-sex marriage made by the members of the Canadian Parliament from both periods. The corpus described here consists of two sets of debate speeches. The first set consisted of 136 *sentences* of the debate speeches and the second set consisted of 400 *paragraphs* of the debate speeches with an average of 70 words. We asked three annotators to examine the statements in the first set with respect to the position of the speaker towards same-sex marriage, and assign *pro*, *con*, or *no* stance. We further asked them to examine which of the pre-existing frames (described in Sect. 3.1) support the statements, and manually annotate them with one of the frames or none; Table 4 shows a few examples from our corpus. To measure inter-annotator agreement, we adopted Weighted Kappa metric. Table 2 shows the achieved agreement for both stance and frames. For almost 90% of the statements, at least two annotators were in agreement. These statements were kept as the final dataset. Some

Table 2. Inter-annotator agreement on parliamentary discourse corpus

	Sentences	Paragraphs
Stance	0.54	-
Frame	0.46	0.70

Table 3. Corpus statistics

Frame	ComArg annotations		Parliamentary annotations	
	Explicit	Implicit	Sentences	Paragraphs
1	16	18	14	16
2	12	18	1	14
3	1	4	0	37
4	50	81	33	55
5	28	52	10	56
6	13	17	2	2
7	56	84	27	63
None	0	0	34	123

statements cannot be judged without their context, and annotators did not agree on the stance or the frame. After discarding the statements for which the annotators were not in agreement, the final set has 121 statements. 87 of these remaining statements are supported by one of the ComArg pre-existing frames.

Unlike the first set, for the paragraph set, we asked the annotators to examine the speeches with respect to only the ComArg frames and ignore the stance. The annotation task for this set was carried out by two annotators, and to check the

Table 4. Examples of frame and stance annotations from parliamentary discourse corpus

Stance	Frame	Parliamentary statements
Pro	4	In my opinion, the answer is clear and simple: two people who want to live together within a civil marriage, regardless of their sexual orientation, must be able to do so without the interference of the State.
Con	7	I urge all members who have even the slightest idea that they want to maintain the definition of marriage that we have known and understood for so long to vote in favour of this so that the government can act on it.
–	1	Let me give an example. When we put something in a category, we are discriminating against everything else that is not in that category. If we have a category of things that are blue, then we are leaving out all the yellows, but that does not mean that blue is better or worse than yellow. It just means that they are different.
–	2	If someone puts a lot into a relationship, into a couple, if someone invests in a house and property, that property has to be protected and we must ensure that if both of them invested, both of them reap the benefits. If one of them dies, at a minimum the inheritance must go to the other or be handled in accordance with the person’s wishes. It should not be possible to deprive someone of what he or she has built up over the years along with his or her spouse. That is not all. There is not only the legal aspect, of course, but also the emotional aspect. We have to change and progress.

reliability, we computed Weighted Kappa (Table 2). The disagreements arose in cases where the speaker used anecdotes or examples. These ambiguous speeches were discarded to create the final dataset. The statistics of the annotated corpora are presented in Table 3.

4 Methods

Distributed word representations are used efficiently in various NLP tasks including sentiment analysis [11]. Recently, embedding models such as those of Mikolov et al. [8], Wang et al. [13], and Kiros et al. [7] have provided an effective and easy way to employ word and sentence representations. These distributed representations are real-valued vectors that capture semantic and syntactic content of words and sentences. Here, we use word and sentence vector representations to measure the semantic textual similarity (STS) between the statements and the frames. Our models then use these similarity measures as features to predict a frame that supports a given statement. We used word2vec embeddings [8] (300-dimensional vectors) trained on Google news articles, and syntactic embeddings [13] (300-dimensional vectors) trained on the Annotated English Gigaword, to compute sentence vectors, and further compare them to skip-thought sentence vectors (4800-dimensional vectors) [7]. Different composition measures are proposed in literature; one of the simplest measures is additive models (Mitchell and Lapata [10]), where word vectors are added together to represent a phrase or sentence representation. Here, we used additive models with word2vec and syntactic vectors to represent the statements (sentences or paragraphs) and we compared them with more complex composition functions based on neural language models. After computing the sentence vectors, we measured the similarity of the statement vector representation with the frame representation. We computed two similarity scores between statements and frames: (1) the cosine similarity of the two vectors, (2) the similarity score represented by a concatenation of the component-wise product of two vectors and their absolute difference (P&D) [12]. We further studied the impact of adding the stance feature (*pro/con*) to the similarity scores as suggested by Boltužić and Šnajder [1]. In addition to the semantic textual similarity and stance features, we also extracted POS-tags, typed dependencies [4], and distributed representations of the statements. Dependency relation features are extracted using the Stanford parser and they represent relationships between pairs of words. For example, for the sentence *We are abandoning traditional liberalism*, the following triples are extracted:

```
nsubj(abandoning-3, We-1)
aux(abandoning-3, are-2)
root(ROOT-0, abandoning-3)
amod(liberalism-5, traditional-4)
doobj(abandoning-3, liberalism-5)
```

Our supervised model then takes these features as input, and learns to identify the frames. For supervised learning, we use SVM^{light} and $SVM^{multiclass}$ by Joachims.^{5,6}

5 Experiment and Results

In the first experiment, we use the statements from ComArg as a training set and the Canadian parliamentary statements on GM as a test set for our classification task. We first remove the stop-words, and then sum the vector representations of the remaining words in the sentences to compute the sentence vectors. For syntactic embeddings, we only used the noun, adjective, and verb embeddings. In case the vector representation for a given word is not found in the embeddings, the lemma of the word is searched and retrieved.

After representing the statements and frames using word2vec, the syntactic-based embedding model, and the skip-thought model, we computed the semantic similarity of each pair with the similarity measures described in Sect. 4.

Our baselines are the majority class and bag-of-words (with TF-IDF vectors and rare words removed) classifiers. Table 5 summarizes our results. We observe that almost all models that use STS features outperform the baselines. We also observe that the P&D similarity score provides a better measure for capturing the meaning of the statement-frame pairs. Furthermore, adding the stance feature to the cosine similarity scores improves the accuracy of the classifiers; however, adding it to P&D has no impact on the accuracy of the classifiers. Although the training set of explicit statements is smaller than the training set of explicit and implicit statements, the best results are mostly achieved by training the classifier on explicit instances. Furthermore, adding the stance feature to the cosine similarity scores gives an improvement of about 20 to 40 percentage points in accuracy above the baseline.

Without using the stance feature, the best score was obtained by training the classifier on explicit and implicit instances with the P&D similarity score of word2vec vectors. While we expected to achieve better accuracy with injecting syntactic information through syntactic embeddings and skip-thought vectors, the results do not show such improvements. This can be due to multiple reasons. First, syntactic embeddings were trained on a smaller set compared to word2vec embeddings. Furthermore, we only rely on three categories of syntactic embeddings (nouns, verbs, and adjectives), whereas even prepositions, such as *against* and *between* are informative features for predicting some frames. Skip-thought models are moreover trained on a dataset with a different genre. One of the challenges of using forum posts as a training set is that they are filled with spelling errors, and in our experiments, we did not correct any of these errors. For our paragraph corpus, since our training corpus based on ComArg is very small, we focused on the two dominant frames in ComArg corpus, frames 4 and 7, and used both explicit and implicit statements for training our models. The

⁵ <http://svmlight.joachims.org/>.

⁶ https://www.cs.cornell.edu/people/tj/svm_light/svm_multiclass.html.

Table 5. Frame prediction results on parliamentary sentences

Train	Features	Accuracy (%)
-	Majority class (argument 4)	33.0
ComArg (Explicit + Implicit)	Bag of words (BoW)	48.2
ComArg (Explicit + Implicit)	STS (Sum of vectors, word2vec, cosine)	54.0
ComArg (Explicit + Implicit)	STS (Sum of vectors, word2vec, cosine) + stance	72.4
ComArg (Explicit + Implicit)	STS (Sum of vectors, word2vec, P&D)	58.6
ComArg (Explicit + Implicit)	STS (Sum of vectors, word2vec, P&D) + stance	58.6
ComArg (Explicit + Implicit)	STS (Sum of vectors, syntactic embeddings, cosine)	49.4
ComArg (Explicit + Implicit)	STS (Sum of vectors, syntactic embeddings, cosine) + stance	68.9
ComArg (Explicit + Implicit)	STS (Sum of vectors, syntactic embeddings, P&D)	50.5
ComArg (Explicit + Implicit)	STS (Skip-thought vectors, cosine)	48.2
ComArg (Explicit + Implicit)	STS (Skip-thought vectors, cosine) + stance	68.9
ComArg (Explicit + Implicit)	STS (Skip-thought vectors, P&D)	51.7
ComArg (Explicit)	Bag of words (BoW)	52.8
ComArg (Explicit)	STS (Sum of vectors, word2vec, cosine)	55.1
ComArg (Explicit)	STS (Sum of vectors, word2vec, cosine) + stance	73.5
ComArg (Explicit)	STS (Sum of vectors, word2vec, P&D)	57.4
ComArg (Explicit)	STS (Sum of vectors, word2vec, P&D) + stance	57.4
ComArg (Explicit)	STS (Sum of vectors, syntactic embeddings, cosine)	54.0
ComArg (Explicit)	STS (Sum of vectors, syntactic embeddings, cosine) + stance	68.9
ComArg (Explicit)	STS (Sum of vectors, syntactic embeddings, P&D)	56.3
ComArg (Explicit)	STS (Skip-thought vectors, cosine)	52.8
ComArg (Explicit)	STS (Skip-thought vectors, cosine) + stance	68.9
ComArg (Explicit)	STS (Skip-thought vectors, P&D)	57.4

paragraph vectors were constructed by adding sentence vectors. For this set, in addition to STS features, we explored features based on POS-tags, the typed dependencies, and the vector representation of the statements. Despite the usefulness of the stance feature as we have seen in the first set, we decided to ignore this feature for our second experiment. The reason for this is that we believe some frames can be used with either positions, for example:

Example 2. *Earlier this year France rejected the marriage of same sex couples because of the effect that same sex marriages have on children.*

Example 3. *Are we going to divide this country into those children who are children of certain couples and children who are not? If we truly value children in the House, then we must understand, as one of the members spoke about children, that this is about the rights of the child, regardless of what their parents do, do not do or who they are.*

Both examples are supported by the frame, *impact on children*; however, the position of the speaker in the first example is against gay marriages, whereas the second speaker supports them.

Similar to the first set, most of the models using STS features outperform the baselines in the paragraph corpus (shown in Table 6). The best results were achieved by the P&D similarity score of word2vec features, followed by the word2vec features extracted from the statements, and typed dependencies. Another observation is that the models based on features extracted from the statements perform better than the models based on cosine similarity features.

Table 6. Frame prediction results on debate paragraph corpus using ComArg corpus (Explicit + Implicit)

Features	Accuracy (%)
Majority class (argument 7)	53.3
Bag of words (BoW)	71.0
Dependency features	72.0
Sum of vectors, word2vec	72.9
Sum of vectors, syntactic embeddings	64.4
STS (Sum of vectors, cosine, word2vec)	61.8
STS (Sum of vectors, P&D, word2vec)	75.4
STS (Sum of vectors, cosine, syntactic embeddings)	61.4
STS (Sum of vectors, P&D, syntactic embeddings)	62.7
STS (Skip-thought vectors, cosine)	53.3
STS (Skip-thought vectors, P&D)	59.3

Table 7. Five-fold cross-validation (4 frames)

Features	Accuracy (%)
Majority class (argument 7)	29.8
Bag of words (BoW)	65.0
POS tags	63.0
Dependency features	53.8
Sum of vectors, word2vec	70.4
Dependency features + word2vec	69.0
Sum of vectors, syntactic embeddings	62.8
Sum of vectors, skip-thought	54.7
STS (Sum of vectors, cosine, word2vec)	42.3
STS (Sum of vectors, P&D, word2vec)	67.6
STS (Sum of vectors, cosine, syntactic embeddings)	39.7
STS (Sum of vectors, P&D, syntactic embeddings)	60.9
STS (Skip-thought vectors, cosine)	41.8
STS (Skip-thought vectors, P&D)	58.6

We further report our results on five-fold cross-validation of four frames (frames 3, 4, 5, and 7) in our paragraph corpus (shown in Table 7). The best results were achieved by the model based on features extracted from the statements, followed by the P&D similarity measure. BOW achieves better performance compared to the other models.

Since the members of the parliament usually refer to the opposing viewpoints and their frames during the debates, relying on all the statements in the

paragraphs for extracting features for the models causes errors. The following example was not successfully tagged with the frame due to treating all the statements in the paragraph in the same way.

Example 4. *Peace River constituents are not opposed to equal rights. In fact, the majority support the legal extension of rights and benefits to same sex couples. However, most are opposed to changing the historical term ‘marriage’ to include these unions. Many have strongly held religious views and are extremely worried that their long-held beliefs are being threatened by the same-sex marriage act. I do not think these views are limited to my riding; I believe they are shared by a majority of Canadians.*

6 Conclusion and Future Work

In this preliminary study, we examined recognizing frames in political argumentative discourse. Many directions have yet to be explored, including (1) discovering frames for various issues, (2) devising a method to deal with larger texts (3) exploring semi-supervised or unsupervised approaches due to the scarcity of human-annotated data for supervised approaches.

Acknowledgements. This work is supported by the Natural Sciences and Engineering Research Council of Canada and by the Social Sciences and Humanities Research Council. We thank Patricia Araujo Thaine, Krish Perumal, and Sara Scharf for their contributions to the annotation of parliamentary statements, and Tong Wang for sharing the syntactic embeddings. We also thank Tong Wang and Ryan Kiros for fruitful discussions, and Christopher Cochrane for insightful comments.

References

1. Boltužić, F., Šnajder, J.: Back up your stance: recognizing arguments in online discussions. In: Proceedings of the First Workshop on Argumentation Mining, pp. 49–58 (2014)
2. Boltužić, F., Šnajder, J.: Identifying prominent arguments in online debates using semantic textual similarity. In: Proceedings of the 2nd Workshop on Argumentation Mining, pp. 110–115. Association for Computational Linguistics, Denver, June 2015
3. Cabrio, E., Villata, S.: Combining textual entailment and argumentation theory for supporting online debates interactions. In: Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers, pp. 208–212. Association for Computational Linguistics (2012)
4. De Marneffe, M.C., Manning, C.D.: The stanford typed dependencies representation. In: Proceedings of the Workshop on Cross-Framework and Cross-Domain Parser Evaluation, COLING 2008, pp. 1–8. Association for Computational Linguistics (2008)
5. Entman, R.M.: Framing: toward clarification of a fractured paradigm. *J. Commun.* **43**(4), 51–58 (1993)

6. Hasan, K.S., Ng, V.: Why are you taking this stance? Identifying and classifying reasons in ideological debates. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 751–762. Association for Computational Linguistics, Doha, October 2014
7. Kiros, R., Zhu, Y., Salakhutdinov, R., Zemel, R.S., Torralba, A., Urtasun, R., Fidler, S.: Skip-thought vectors. arXiv preprint [arXiv:1506.06726](https://arxiv.org/abs/1506.06726) (2015)
8. Mikolov, T., Chen, K., Corrado, G., Dean, J.: Efficient estimation of word representations in vector space. arXiv preprint [arXiv:1301.3781](https://arxiv.org/abs/1301.3781) (2013)
9. Misra, A., Anand, P., Tree, J.E.F., Walker, M.A.: Using summarization to discover argument facets in online ideological [sic] dialog. In: The 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL HLT 2015, Denver, Colorado, USA, May 31–June 5, 2015, pp. 430–440 (2015)
10. Mitchell, J., Lapata, M.: Vector-based models of semantic composition. In: Association for Computational Linguistics, pp. 236–244 (2008)
11. Socher, R., Pennington, J., Huang, E.H., Ng, A.Y., Manning, C.D.: Semi-supervised recursive autoencoders for predicting sentiment distributions. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing, pp. 151–161. Association for Computational Linguistics (2011)
12. Tai, K.S., Socher, R., Manning, C.D.: Improved semantic representations from tree-structured long short-term memory networks. arXiv preprint [arXiv:1503.00075](https://arxiv.org/abs/1503.00075) (2015)
13. Wang, T., Mohamed, A., Hirst, G.: Learning lexical embeddings with syntactic and lexicographic knowledge. In: Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (vol. 2: Short Papers), pp. 458–463. Association for Computational Linguistics, Beijing, July 2015



<http://www.springer.com/978-3-319-46217-2>

Principles and Practice of Multi-Agent Systems
International Workshops: IWEC 2014, Gold Coast, QLD,
Australia, December 1-5, 2014, and CMNA XV and IWEC
2015, Bertinoro, Italy, October 26, 2015, Revised
Selected Papers

Baldoni, M.; Baroglio, C.; Bex, F.; Grasso, F.; Green, N.;
Namazi-Rad, M.-R.; Numao, M.; Suarez, M.T. (Eds.)

2016, X, 169 p. 39 illus., Softcover

ISBN: 978-3-319-46217-2