



University of Dundee

A Corpus of Argument Networks

Konat, Barbara; Lawrence, John; Park, Joonsuk; Budzynska, Katarzyna; Reed, Chris

Published in:
LREC 2016

Publication date:
2016

Document Version
Final published version

[Link to publication in Discovery Research Portal](#)

Citation for published version (APA):

Konat, B., Lawrence, J., Park, J., Budzynska, K., & Reed, C. (2016). A Corpus of Argument Networks: Using Graph Properties to Analyse Divisive Issues. In N. Calzolari, K. Choukri, T. Declerck, S. Goggi, M. Grobelnik, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odiijk, ... S. Piperidis (Eds.), LREC 2016: Proceedings for the Tenth International Conference on Language Resources and Evaluation. (pp. 3899-3906). Slovenia.

General rights

Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from Discovery Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

A Corpus of Argument Networks: Using Graph Properties to Analyse Divisive Issues

Barbara Konat¹, John Lawrence¹, Joonsuk Park²,
Katarzyna Budzynska^{3,1}, Chris Reed¹

¹Centre for Argument Technology, University of Dundee, Scotland

²Department of Computer Science, Cornell University, New York, USA

³Institute of Philosophy and Sociology, Polish Academy of Sciences, Poland

Abstract

Governments are increasingly utilising online platforms in order to engage with, and ascertain the opinions of, their citizens. Whilst policy makers could potentially benefit from such enormous feedback from society, they first face the challenge of making sense out of the large volumes of data produced. This creates a demand for tools and technologies which will enable governments to quickly and thoroughly digest the points being made and to respond accordingly. By determining the argumentative and dialogical structures contained within a debate, we are able to determine the issues which are divisive and those which attract agreement. This paper proposes a method of graph-based analytics which uses properties of graphs representing networks of arguments pro- & con- in order to automatically analyse issues which divide citizens about new regulations. By future application of the most recent advances in argument mining, the results reported here will have a chance to scale up to enable sense-making of the vast amount of feedback received from citizens on directions that policy should take.

Keywords: Argument Analysis, Argument Structure, Automatic Text Interpretation, Divisiveness Measures, e-Participation, Graph-based Analytics

1. Introduction

The aim of this paper is to show how the analysis of argumentative and dialogical structures, and in particular pro- and con-arguments, allows for the principled identification of *divisive* issues in an online corpus of debates. More precisely, we show that structuring data as networks of interacting arguments, claims, reasons and conflicts means that the data is represented in the form of a graph. This representation, in turn, offers the possibility of using graph properties to automatically analyse which issues attracted most attention and which most divided disputants.

We present an annotated corpus which builds on previous work in argument mining (Park and Cardie, 2014) by adding further mark-up capturing dialogical interactions between participants in an eRulemaking system for online deliberative democracy. We demonstrate how the rich network structure available in this style of annotation supports multiple interpretations of divisiveness which in turn provide powerful insights into the nature of the debate – insights which can be valuable in creating summaries and overviews of complex, multi-faceted debates for decision-makers who need sense-making tools to develop a coherent understanding of large volumes of data from public contributions – in this case on contentious policy issues within the remit of the Department of Transportation in the US.

The paper is structured as follows. Section 2 discusses some related work in argument analysis and controversy detection. Section 3 presents language resources used in this study. Section 4 demonstrates the type of information which can be extracted by applying our model – we show how the structure of argument networks can be used to calculate measures of divisiveness with respect to citizens' opinions about new regulations. Finally, Section 5 shows the automation of the model.

2. Related work

Automated identification of divisive issues in the context of argumentative dialogue is closely related to two research areas: argument analysis and controversy detection.

2.1. Argument analysis

Argument analysis is one of two key areas, together with argument evaluation, of the theory of argumentation (see e.g. (van Eemeren et al., 2014)). The most recent advances in computational linguistics have been exploring methods and techniques for making automation of this process possible and efficient. Argument mining aims at extraction of argument structures from natural language texts (see e.g. (Peldszus and Stede, 2013)). Models of arguments adopted in computational approaches vary with regards to their complexity. In some studies the argument is understood as just one proposition, whereas in others as a binary, reason-conclusion pair. The theory of argumentation however allows for more developed and complex conceptualisations, using argument diagrams (Toulmin, 1958), argument schemes (Walton et al., 2008), distinguishing between convergent and linked arguments (Freeman, 1991) or between supports and attacks (Dung, 1995). The initial stage in argument mining consists of manual annotation of an argument dataset, using tools which implement the theoretical concepts. Several resources and tools for manual annotation and visualisation of arguments are available, including Carneades (Gordon et al., 2007), AIFdb (Lawrence et al., 2012) and Araucaria (Reed and Rowe, 2004)).

In the last decade, attempts have been made to automate the process of argument analysis, as manual annotation is very time consuming. Text genres used in the studies range from restricted, formalised language (legal texts or academic papers) to more unstructured natural language (such as spoken dialogue or Internet fora). The field of argument mining started to gain interest over a decade ago with Ar-

gumentative Zoning - an attempt at automated extraction of arguments from scientific papers (Teufel, 1999; Teufel and Moens, 2002). In the study, arguments are understood as spans of texts serving various argumentative functions, and automated recognition obtained the highest F-score of 0.86 for the recognition of parts of papers in which an author refers to their own research and the lowest F-score of 0.26 for the recognition of parts in which an author presents arguments against other approaches. In (Moens et al., 2007) similarly, spans of texts (sentences) are classified as either argumentative or non-argumentative, finding that automated extraction of arguments is more efficient from well-structured texts than from informal Internet debates (for example newspaper articles achieved 73.22% accuracy whereas discussion fora – 68.4%). Mining of arguments understood as graphs, explored in (Palau and Moens, 2009) results in 73% accuracy on the Araucaria corpus and 80% on the ECHR corpus (a corpus of texts issued by the European Court of Human Rights). Another set of well-structured argumentative texts (“microtexts” in German) is used in (Peldszus, 2014) and provides a highest achieved F-score of 0.7 for automated extraction of reason-conclusion structures.

Mining the arguments from less structured texts can be supported with various methods. Combining the statistical approach with the application of argument schemes to argument mining (Lawrence and Reed, 2015) allows for the F-score of 0.83 on the corpus of Internet arguments. Naturally occurring dialogues (either face-to-face or on-line) do not have pre-defined structure, which makes extraction of arguments even more difficult. This is why mining arguments from dialogue (Budzynska et al., 20xx) explores how argumentative structures are built upon dialogical structures (mainly illocutionary connections) in interaction.

2.2. Controversy detection

Controversy detection seeks to develop automated methods for identifying events or issues which attract conflicting opinions. Controversy of a given issue can be measured by the number of revisions of an article in Wikipedia (Kittur et al., 2007), or the number of Tweets concerning a certain topic over a time window (Popescu and Pennacchiotti, 2010). Controversy can be also detected using sentiment analysis (Choi et al., 2010), when conflicting sentiments expressed towards a given issue by different users is understood as an indicator of controversy. In Internet Argument Corpus (Walker et al., 2012), agreement and disagreement in dialogue are treated as indicators of controversy (Rosenthal and McKeown, 2015). Extracting controversy in the Web can also enable users to be informed of controversial issues and alerted when alternative viewpoints are available, as presented in (Dori-Hacohen and Allan, 2013). Analysing online ideological dialogues in (Misra et al., 2015) connects the concept of argument and controversy, by studying controversy in the context of central claim and the concept of argument facet.

There are two most distinctive features that distinguish these works from ours. First, our goal is not to mine controversial issues directly, but to mine arguments pro- & con- to create argument maps as graphs, and then use graph prop-

erties to automatically compute controversial issues. This means our work presented here constitutes only one (final) stage of the argument mining process, i.e. the automated interpretation and summarization of mined data. Second, automatic extraction of just two categories (arguments pro- & con-) allows for the more fine-grained overview of comments on policies. Using the apparatus of graph theory, we can introduce different and precise definitions of divisiveness (see Section 4.2. for two such definitions) in order to capture various interpretations of what divides people. In order to make a clear distinction between our approach and that of controversy mining, we introduce a distinctive term ‘divisive’ instead of ‘controversial’.

3. Material and methods

3.1. e-Participation in the US

eRulemaking (online deliberative democracy) is a multi-step process of social media outreach that US federal agencies use to consult with citizens about new regulations on health and safety, finance, and other complex topics. In order to ensure public awareness and participation of new regulations, agencies are required to publish materials describing the legal basis, factual and technical support, policy rationale, and costs and benefits of a proposal. They must specify a comment period, usually 60 to 90 days, during which anyone may send the agency their comments. Further, agencies are required to respond to information, arguments, and criticisms presented by the public as part of its final rule (Lubbers, 2012).

Once the agency introduces the new regulation (either in the original form or a modified one as a result of public consultation), it is obliged to summarise the comments it received; respond to questions and criticisms; and explain why it did not make changes. In case of complaints from citizens, the court will use this documentation in order to make sure that the agency respected public comments in a satisfactory way. The court will return the regulation to the agency for modification, if it decided that some citizens’ suggestions or criticism were not addressed.

RegulationRoom.org is one of a number of eRule-making platforms which host regulation proposals from various US government agencies allowing people to submit online comments. Such platforms allow for collecting a large amount of socially valuable data (e.g. in the US over 200 million citizens are eligible to vote and thus can participate in RegulationRoom (Farina and Newhart, 2013; Park et al., 2012)). On the other hand the volume of data makes the task of its interpretation and summarisation extremely challenging. This work aims to address this problem.

3.2. Datasets

Our corpus is comprised of user comments extracted from RegulationRoom.org. First, we transferred part of the annotated data from *Airline Passenger Rights* (APR) rule – a subset of a yet unpublished corpus collected at Cornell containing the relation labels of pro-arguments.¹ The

¹*Airline Passenger Rights* is one of the several rules comprising the corpus. See Park et al. (2015) for the descriptions of the pro- relations.

APR-Cornell dataset consists of 923 comments and 8,320 propositions (segments).

In the next step, we selected only those comments which had dialogical nature, i.e. which attracted at least one reply. This dataset was called Regulation Room Divisiveness (RRD). It consists of 209 comments, which in this case constitute turns in the dialogical exchange, and 70 maps which are graphs representing argument networks. The annotation was extended by adding more pro-arguments (using more fine-grained criteria) and con-arguments which are inherently dialogical (see Table 1 for the size of language resources used in this study).

Table 1: Summary of the language resources: Cornell corpus for *Airline Passenger Rights* (APR) discussion; and the Regulation Room Divisiveness (RRD) corpus.

	Words	Segments	Comments	Maps
APR	118,789	8,320	923	-
RRD	23,682	1,657	209	70

3.3. Regulation Room Divisiveness Corpus

The annotation was performed using the OVA+ annotation tool (Janier et al., 2014)² marking three types of relations between propositional contents of comments (see Table 2): pro-arguments (Default Inference, RA); con-arguments (Default Conflict, CA); and the relation of Rephrase (Default Rephrase, MA), which captures situations when people give the same comment, but use a different linguistic surface.

In the annotation, the argumentative function is understood as the relation between two propositions, not as the property of one span of text. In the OVA+ tool, these relations are marked as edges connecting information nodes (I-nodes) which contain propositions (see Fig. 1). To convert the raw text into an argument map, the analyst needs to paste the text into the left hand panel and then click on the right hand panel to create an I-node. Edges can be created by clicking the “Add edge” button and dragging the mouse between I-nodes. After the annotation, the map can be saved to the AIFdb database³ (Lawrence et al., 2012) and later downloaded in various file formats (including .json and .pl). The tool is a web-based application, freely available to use for annotation of argument diagrams.

Default Inference holds between two propositions when one proposition provides a reason to accept another proposition. In other words, a supporting claim can be used to answer the question “why *p*?”. In the example (1) from the map RRD:#4900, (1-a) provides support for (1-b). If the propositional content of (1-a) was challenged in a dialogical situation with the question “why?”, proposition (1-b) could be used as an answer to this question. In this example the user “SBARB95” is arguing that the suggested regulation (obligating airlines to inform passengers about delays longer than 30 minutes) should not be introduced. As the reason for this claim, the user “SBARB95” provides

the statement that it usually takes longer than 30 minutes to travel to the airport.

- (1) a. SBARB95: *In my experience it usually takes about 30 minutes to get to major airports*
- b. SBARB95: *I wonder if delays of 30 minutes would actually affect passenger behavior*

Default Conflict holds between two propositions which cannot be both true at the same time. Speakers use conflicting propositions to attack another speaker’s claims, by means of providing counter-claims. Example (2) from the map RRD:#4891 presents the situation in which the claim (2-a) provided by one user is attacked with the claim (2-b) by another user. In the example, user “AKTRAVELLER” suggests a new regulation, according to which the airlines should inform passengers in advance about possible delays or cancellations. This statement is attacked by the user “SOFIEM”, who is providing a counter-claim, saying that the solution is not possible.

- (2) a. AKTRAVELLER: *The airline could call in advance and give the passenger their options*
- b. SOFIEM: *Unfortunately, there’s no way to give advance notice*

Default Rephrase holds between two propositions with the same or similar content expressed with different linguistic surface. Our concept of Rephrase is quite broad and covers all propositions serving the same argumentative function (e.g. repeated conclusions or premises) even in cases where the meaning equivalence of the propositions is not complete. We decided to annotate the relation of Rephrase to capture the fact that rephrased content does not constitute additional support (i.e. one propositional content repeated three times should not be counted as three separate supports for a claim). In the example (3) from the map RRD:#5411) one speaker is rephrasing their own conclusion (3-a) by restating similar propositional content in (3-b). The user “DBERGER” repeats and reformulates their opinion concerning regulation on peanuts being consumed on the planes.

- (3) a. DBERGER: *There must be a complete ban on tree nuts and peanuts on planes*
- b. DBERGER: *Again all nuts should be banned from airplanes*

These binary annotations of relations between propositions create the “building blocks” of argument networks. Results of simple annotations of examples (1), (2), (3) are presented in Fig. 2, and Section 4. describes in more detail how they are used in creation of more complex argument networks.

Table 2 presents a summary of relations of Inference, Conflict and Rephrase in the RRD corpus. Regulation Room Divisiveness corpus is freely available at <http://arg.tech/rrd>. The corpus uses the open Argument Interchange Format (AIF) standard for argument representation (Rahwan et al., 2007) and constitutes a part of the AIFdb database.

²Available at <http://ova.arg-tech.org>

³Available at <http://aifdb.org>

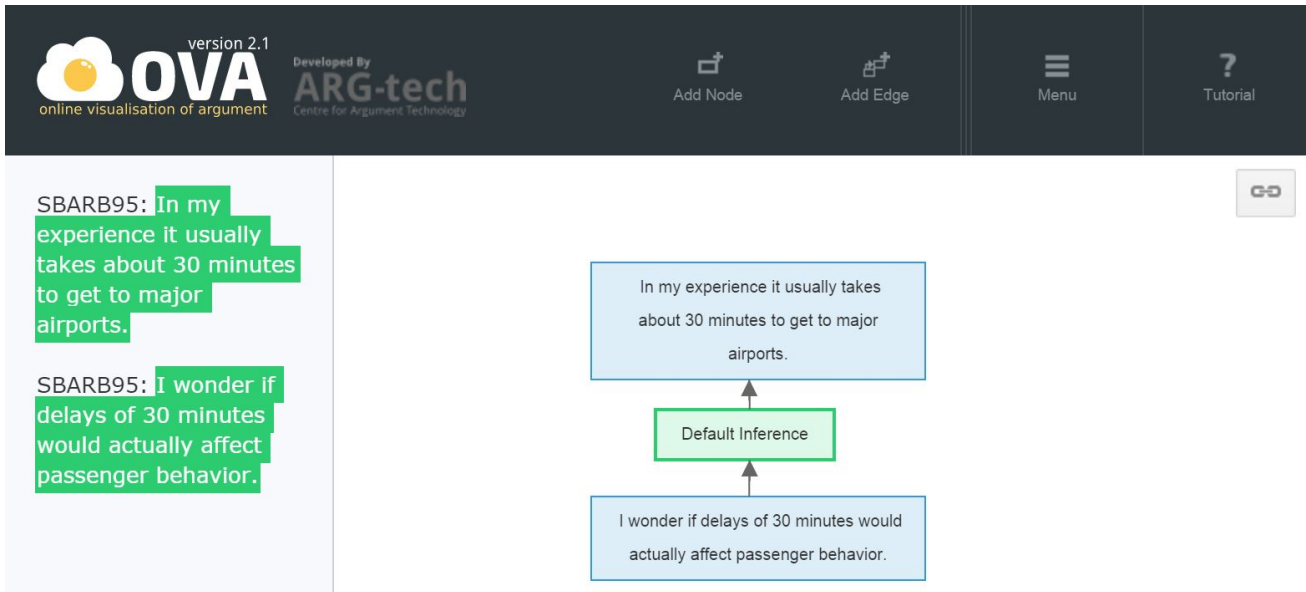


Figure 1: OVA+ – Online Visualisation of Arguments tool: annotation window.

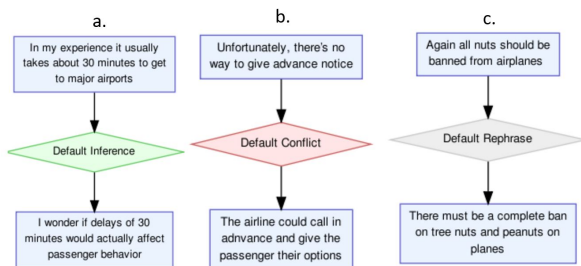


Figure 2: Relations of Inference, Conflict and Rephrase as building blocks of argument networks.

Table 2: Occurrences of relations between contents of comments in the annotated corpus of argument networks.

Relation type	Number
Inference (RA)	671
Conflict (CA)	97
Rephrase (MA)	14
Total	782

4. Divisiveness in argument networks

4.1. Argument network

Spotting quickly the most divisive issues in a large data of texts is a real challenge. Consider a conversation between three RegulationRoom users about whether or not peanuts should be prohibited on planes as they may cause an allergic reaction.

- (4) a. MALLONE: *When a food allergy is life threatening (and known to cause anaphylaxis), it considered a disability under federal laws such as Section 504 of the Rehabilitation Act of 1973 and the Americans with Disabilities Act (ADA).*
- b. *In other words, people with severe peanut allergies have the right to be protected.*

- c. MULDER: *No, allergies are not disabilities, and therefore you get no special treatment under the ADA.*
- d. *Federal courts have consistently ruled this way.(..)*
- f. ANTANAGOGE: *Mulder's comment about the ADA is only partially true, but thoroughly exaggerated. because there has only been one court case.*
- h. *Food allergy is generally considered a disability under Section 504 and ADA.*
- i. *The point Mulder exaggerates is that there is no primary legal precedent, i.e., a court opinion, saying this.(...)*
- j. *The ACAA prohibits discrimination against those with disabilities by U.S. and foreign air carriers, and DOT regulations require airlines to accommodate travelers with disabilities.*

Even in such a short excerpt of online comments, it is not trivial to identify the most divisive issue. However the task becomes much easier if we unpack the structure of this dialogue and create its representation as a directed graph of pro-arguments (green nodes) and con-arguments (red nodes) (see Fig. 3; manually annotated using the OVA+ tool). Notice that it is now clear that the issue (4-d) attracted the highest number of arguments, even though it is (4-b) which is a main claim of the discussion.

Divisiveness can be conceptualised as a feature of an issue, but also as the strength of conflict between two issues. In Fig. 3 from the total number of five conflict relations, the one between (4-d) and (4-f) attracts the highest number of pro-arguments, i.e. (4-c) for (4-d); and (4-g) and (4-i) for (4-f). Intuitively, the more support two conflicting issues have, the higher strength of the conflict between them.

4.2. Graph properties

In order to obtain operational definitions and provide measurable criteria for extraction of divisive issues we founded our concepts on the graph properties. In the corpus of argument network, divisiveness of a given proposition can be

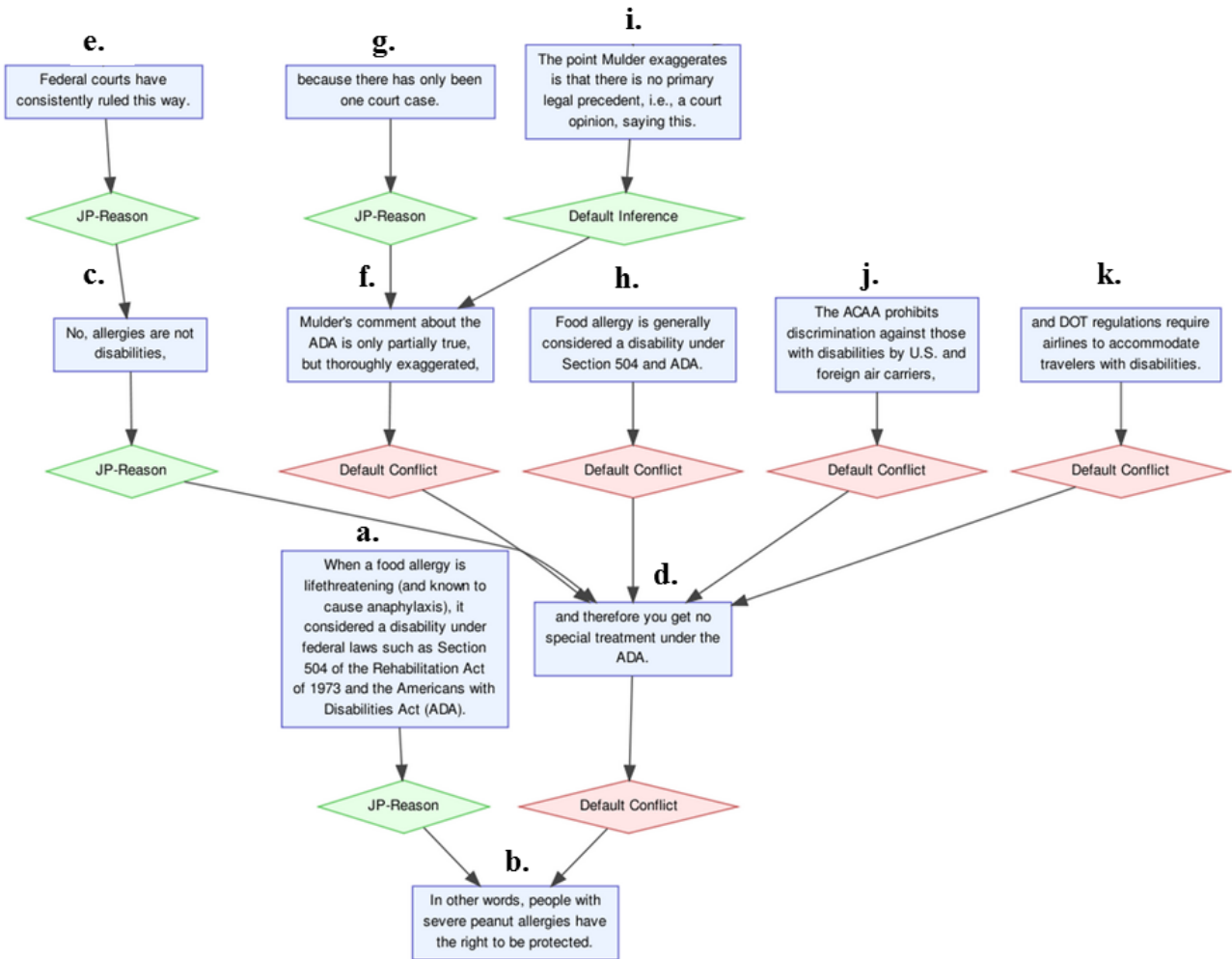


Figure 3: Fragment of the argument network with the example of divisive issues (RRD:# 5695). Full map available in the Regulation Room Divisiveness corpus at <http://arg.tech/rrd>.

defined with the properties of the graph.

We construe the argumentation analysis as a directed graph, $G = (V, E)$, in which vertices (V) are either propositions or relations between propositions, and those relations are either support (pro-arguments) or conflict (con-arguments), captured by a function R which maps $V \mapsto \{prop, support, conflict\}$ and edges exist between them $E \subset V \times V$. For syntactic convenience, we refer to the number of edges (i.e. the IN-order of) at a vertex v as $|v|$ and add superscript to indicate whether we are interested in the number of incoming or outgoing edges, and a superscript to indicate whether the edge connects v with a support or a conflict, e.g., $|v|_{support}^{in}$.

A first estimate of the divisiveness of a proposition v , which captures the intuition behind assessing the issue (4-d) in Example (4) as the most divisive, can be calculated thus:

$$D_1(v) = |v|_{support}^{in} * |v|_{conflict}^{in}$$

Alternatively, the divisive issue may be described with the strength of the conflict to capture the interpretation of the divisiveness of the conflict between (4-d) and (4-f). In this case it will be measured as the relation between two nodes v_1, v_2 which are in conflict. For that, we need to account for the number of supports provided for each of the nodes.

Thus, given $(v_1, v_c) \in E$ and $(v_c, v_2) \in E$ and $R(v_c) = conflict$:

$$D_2(v_1, v_2) = |v_1|_{support}^{in} * |v_2|_{support}^{in}$$

The two concepts of divisiveness presented here allow for automated identification of divisive issues in the corpus of argument networks.

5. Results

These two measures were implemented as a first step of the development of an Graph-based Analytics tool available at <http://arg.tech/rrdgraph>. Fig. 4 and 5 show the automatically identified divisiveness scores, D_1 and D_2 respectively, for our corpus as a whole. In each chart, the score is plotted against the number of propositions with that score. For example, in Fig. 4, we can see that there are twenty propositions with a D_1 score of 1 (i.e. with one incoming conflict and one incoming support), nine with a D_1 score of 2, four with a score of 3, four with a score of 4, and one proposition with a D_1 score of 6. The four propositions with D_1 score of 4 are listed below:

1. I fully support a ban on peanuts and food containing peanuts [4 support, 1 conflict]

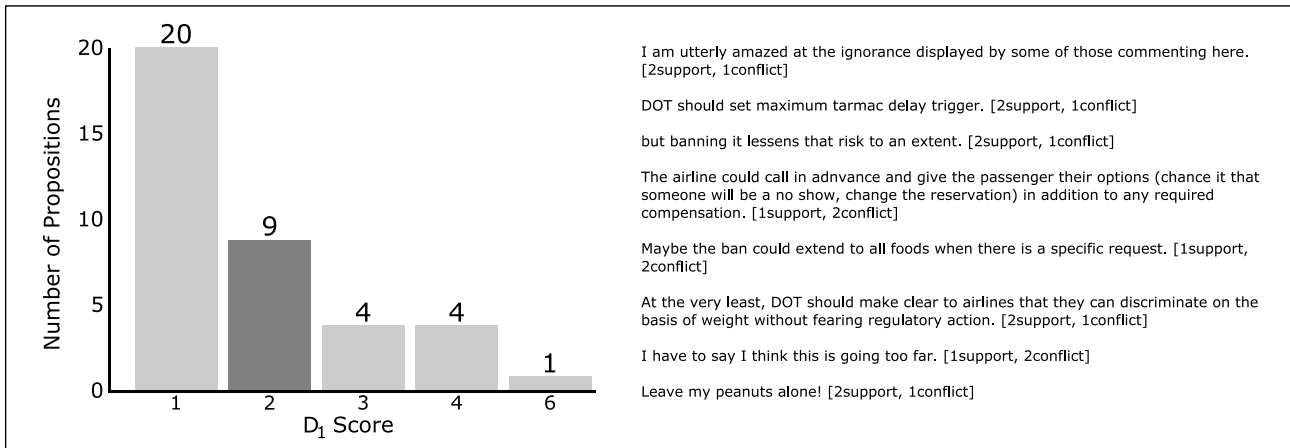


Figure 4: Chart showing Divisiveness Score D_1 . The text on the right hand side lists a selection of the issues with a D_1 score of 2 (highlighted in darker grey), along with the number of supporting and conflicting items each issue has.

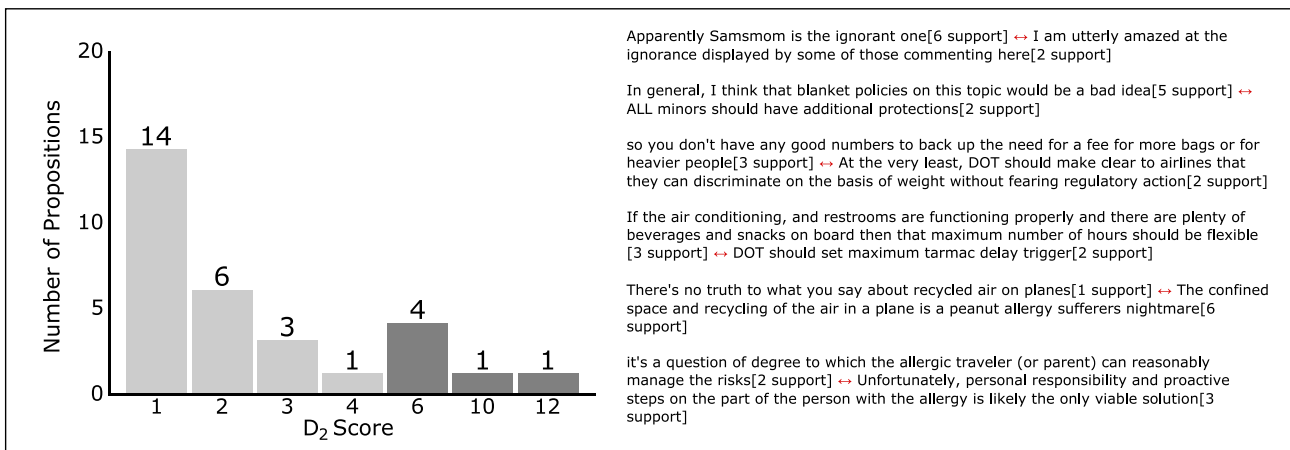


Figure 5: Chart showing Divisiveness Score D_2 . The text shows those pairs of conflicting issues with the highest D_2 scores (of 6, 10 and 12), along with the number of supporting items for the issue on each side.

2. ALL minors should have additional protections [2 support, 2 conflict]
3. I would support a full ban of peanut products on any airline [4 support, 1 conflict]
4. and therefore you get no special treatment under the ADA [1 support, 4 conflict]

Notice that the system automatically spotted the divisive issue (4-d) in the argument network depicted in Fig. 3: (4-d) is the fourth proposition listed above.

Fig. 5 shows that there is one conflict identified which especially divides people, with a D_2 score of 12. This is the conflict between the propositions “Apparently Samsmom is the ignorant one” and “I am utterly amazed at the ignorance displayed by some of the users here”, with the first of these propositions having six supporting claims and the second proposition having two supporting claims. The divisive conflict between (4-d) and (4-f) in Fig. 3, has a D_2 score of 2, making it one of the sixteen most divisive conflicts within the corpus.

Validation of these results by comparison to human annotations of divisive issues is, unfortunately, not a straight-

forward task. Divisiveness is not a property intrinsic to a specific issue or comment, but is a result of how different comments interact. With a large volume of data to process, it is impossible for an annotator to process the connections and dependencies between hundreds of comments, and so determine those which are divisive. Additionally, the fact that divisiveness occurs when multiple people are in disagreement, either over a single issue, or are split in favour of two conflicting points makes this a challenging concept for a single annotator to identify: an issue which seems perfectly clear to the annotator may in fact be divisive in the context of a multi-participant dialogue. However, the measures of divisiveness which we present here give scores based precisely on these criteria, meaning that, as long as the dialogue structure has been correctly captured, those issues with the highest D_1 and D_2 will be precisely those that most divided people.

6. Conclusion

We have presented a corpus of user comments on the RegulationRoom.org platform using labels for pro and con-arguments, and rephrase. We have shown that we are able to use the resulting graph structure to gain insights

and understanding of the complex interactions through the Argument Graph Analytics tool. The two measures of divisiveness which we have introduced, whilst simple to calculate, give a clear indication of where contentious issues in the discussion are taking place, and match closely with an intuitive understanding of the topic.

Though the work described in this paper relies upon manual annotation of data, automatic techniques for identifying argument structure could be used as input to these same processes. As these argument mining techniques (Peldszus and Stede, 2013; Lawrence and Reed, 20xx; Budzynska et al., 20xx) continue to improve, the ability to interpret the resulting argument network will be a valuable tool in creating fully automated summaries and overviews of complex, multi-faceted debates, supporting deep insights into large volumes of argument and debate data.

7. Acknowledgments

The authors would like to acknowledge that the work reported in this paper has been supported in the UK in part by RCUK Lifelong Health and Wellbeing Programme grant number EP/K037293/1 - BESiDE: The Built Environment for Social Inclusion in the Digital Economy, in part by the EPSRC grant number EP/N014871/1 Argument Mining and in part by the Innovate UK under grant 101777 and in the US in part by NSF grant IIS-1314778. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation or other funders.

8. References

- Budzynska, K., Janier, M., Kang, J., Konat, B., Reed, C., Saint-Dizier, P., and Yaskorska, O. (20xx). Automatically identifying transitions between locutions in dialogue, in press. In *Proceedings of 1st European Conference on Argumentation: Argumentation and Reasoned Action (ECA 2015)*.
- Choi, Y., Jung, Y., and Myaeng, S.-H. (2010). Identifying controversial issues and their sub-topics in news articles. In *Proceedings of the Intelligence and Security Informatics Conference, Pacific Asia Workshop, PAISI 2010*, pages 140–153.
- Dori-Hacohen, S. and Allan, J. (2013). Detecting controversy on the web. In *Proceedings of the 22nd ACM International Conference on Information & Knowledge Management*, pages 1845–1848. ACM.
- Dung, P. M. (1995). on the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77(2):321–357.
- Farina, C. R. and Newhart, M. J. (2013). Rulemaking 2.0: Understanding and getting better public participation. In *Cornell e-Rulemaking Initiative Publications*.
- Freeman, J. B. (1991). *Dialectics and the macrostructure of arguments: A theory of argument structure*, volume 10. Walter de Gruyter.
- Gordon, T. F., Prakken, H., and Walton, D. (2007). The Carneades model of argument and burden of proof. *Artificial Intelligence*, 171(10):875–896.
- Janier, M., Lawrence, J., and Reed, C. (2014). OVA+: An argument analysis interface. In *Proceedings of the Fifth International Conference on Computational Models of Argument (COMMA 2014)*, pages 463–464.
- Kittur, A., Suh, B., Pendleton, B. A., and Chi, E. H. (2007). He says, she says: Conflict and coordination in Wikipedia. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 453–462. ACM.
- Lawrence, J. and Reed, C. (2015). Combining argument mining techniques. In *Working Notes of the 2nd Argumentation Mining Workshop, ACL'2015*.
- Lawrence, J. and Reed, C. (20xx). Argument mining, in review.
- Lawrence, J., Bex, F., Reed, C., and Snaith, M. (2012). AIFdb: Infrastructure for the Argument Web. In *Proceedings of the 4th International Conference on Computational Models of Argument COMMA*, pages 515–516.
- Lubbers, J. (2012). *A Guide to Federal Agency Rulemaking*. ABA Section of Administrative Law and Regulatory Practice and Government and Public Sector Lawyers Division.
- Misra, A., Anand, P., Tree, J. F., and Walker, M. (2015). Using summarization to discover argument facets in online ideological dialog. In *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics – Human Language Technologies (NAACL HLT 2015)*.
- Moens, M.-F., Boiy, E., Palau, R. M., and Reed, C. (2007). Automatic detection of arguments in legal texts. In *Proceedings of the 11th International Conference on Artificial Intelligence and Law (ICAIL)*, pages 225–230. ACM.
- Palau, R. M. and Moens, M.-F. (2009). Argumentation mining: the detection, classification and structure of arguments in text. In *Proceedings of the 12th International Conference on Artificial Intelligence and Law (ICAIL)*, pages 98–107. ACM.
- Park, J. and Cardie, C. (2014). Assess: A tool for assessing the support structures of arguments in user comments. In *Proceedings of the Fifth International Conference on Computational Models of Argument (COMMA 2014)*, pages 473–474.
- Park, J., Klingel, S., Cardie, C., Newhart, M., Farina, C., and Vallbé, J.-J. (2012). Facilitative moderation for online participation in erulemaking. In *Proceedings of the 13th Annual International Conference on Digital Government Research*, pages 173–182. ACM.
- Park, J., Blake, C., and Cardie, C. (2015). Toward machine-assisted participation in eRulemaking: An argumentation model of evaluability. *Proceedings of the 15th International Conference on Artificial Intelligence and Law (ICAIL)*.
- Peldszus, A. and Stede, M. (2013). From argument diagrams to argumentation mining in texts: a survey. *International Journal of Cognitive Informatics and Natural Intelligence (IJCINI)*, 7(1):1–31.
- Peldszus, A. (2014). Towards segment-based recognition

- of argumentation structure in short texts. In *Proceedings of First Workshop on Argumentation Mining*.
- Popescu, A.-M. and Pennacchiotti, M. (2010). Detecting controversial events from Twitter. In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management (CIKM)*, pages 1873–1876. ACM.
- Rahwan, I., Zablith, F., and Reed, C. (2007). Laying the foundations for a World Wide Argument Web. *Artificial Intelligence*, 171(10-15):897–921.
- Reed, C. and Rowe, G. (2004). Araucaria: Software for argument analysis, diagramming and representation. *International Journal on Artificial Intelligence Tools*, 13(4):961–980.
- Rosenthal, S. and McKeown, K. (2015). I couldn't agree more: The role of conversational structure in agreement and disagreement detection in online discussions. In *16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 168–177.
- Teufel, S. and Moens, M. (2002). Summarizing scientific articles: Experiments with relevance and rhetorical status. *Computational linguistics*, 28(4):409–445.
- Teufel, S. (1999). *Argumentative Zoning: Information Extraction from Scientific Text*. Ph.D. thesis, University of Edinburgh.
- Toulmin, S. E. (1958). *The uses of argument*. Cambridge University Press.
- van Eemeren, F. H., Garssen, B., Krabbe, E. C., Snoeck Henkemans, A., Verheij, B., and Wagemans, J. H. (2014). *Handbook of Argumentation Theory*. Springer.
- Walker, M. A., Tree, J. E. F., Anand, P., Abbott, R., and King, J. (2012). A corpus for research on deliberation and debate. In *Proceedings of 8th International Conference on Language Resources and Evaluation (LREC)*, pages 812–817.
- Walton, D., Reed, C., and Macagno, F. (2008). *Argumentation Schemes*. Cambridge University Press.