Computational Models of Argument
C. Reed et al. (Eds.)

doi:10.3233/FAIA240343

Documents with Integrated Visually Annotated Arguments

Florian RUOSCH [1], Joel WATTER, Cristina SARASUA and Abraham BERNSTEIN
Department of Informatics, University of Zurich, Switzerland

Keywords. Argument Annotations, Argument Mining, Argument Corpora

1. Introduction

Argumentation is crucial in scientific communication [1]. It is used to reason for new ideas and to argue about existing works. Therefore, the underlying argumentative structure in scholarly publications conveys a plethora of information. Argument Mining (AM) [2] intends to automate the process of discovering arguments, their components, and their relations in natural language texts. The format for the extracted arguments has only been standardized to a certain degree and often splits the original running text from the annotations, which makes reading (and understanding) harder for a human. Meanwhile, they are necessitated by current AM systems. Grasping an annotated claim's content (and, possibly, context) is not trivial. An even more challenging task is to compare different annotations from, for example, various AM systems to evaluate which of them is the most useful. It would, hence, be useful to visualize these annotations and put them into their original context in the running text for human readers while keeping the capability to use them for training or evaluation.

To this end, we present the DIVAA tool to create **D**ocuments with **I**ntegrated **V**isually **A**nnotated **A**rguments.[2] Its pipeline converts pre-existing annotations in the `BRAT` standoff format to PDF, a widely used and platform-independent document format. The DIVAA tool embeds the original two `BRAT` files (i.e., the text and the annotations) as attachments in the PDF and also provides a way to export them. This unification makes sharing new annotations easier because they are in a single file.

2. Demonstration

To illustrate the use of the DIVAA tool, we converted the 40 argument-annotated papers of the Sci-Arg corpus [4] to PDFs with visually annotated arguments. Figure 1 shows a screenshot of an annotated PDF file opened in Adobe Acrobat Reader. The leftmost widget, titled `Attachments`, shows the two original annotation files consisting of the .txt containing the full text and the .ann with the `BRAT` standoff format annotations. These

[1]Corresponding Author: Florian Ruosch, ruosch@ifi.uzh.ch.
[2]https://gitlab.ifi.uzh.ch/DDIS-Public/DIVAA

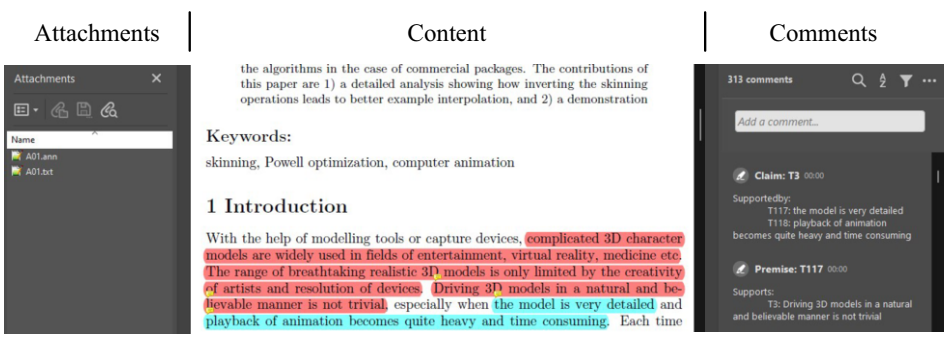

Figure 1. Screenshot of an annotated example file [3] from the Sci-Arg corpus [4].

two can be exported for downstream tasks using the extraction DIVAA module. Even though they appear as separate files, they are attached within the PDF, and therefore, only a single file needs to be shared or distributed. The middle (`Content`) shows the PDF content as formatted text with spans highlighted. These colors correspond to the different components that represent arguments. More information about these annotations is displayed in the last widget. The rightmost part of the interface contains annotations using the `Comments` functionality of the PDF. Each component is converted to exactly one comment containing information about it. The comment title reflects the component type and the identifier from the `BRAT` standoff format. Additionally, the comment body shows other argumentative components to which this one is related and the corresponding relation type. Clicking on a comment takes the reader to the appropriate position in the text where the component is located. The same also holds the other way around: selecting a highlighted part of the text shows the respective comment. These features aim to facilitate navigation and inspection of the annotations by embedding them into the full text.

Acknowledgments. This research was partially funded by the Swiss National Science Foundation (SNSF) through projects "CrowdAlytics" (contract 184994) and "Digital Deliberative Democracy" (contract 205975).

References

[1] Khatib KA, de Waard A, Freitag D, Gurevych I, Hou Y, Scells H. Towards a Unified Model of Scholarly Argumentation (Dagstuhl Seminar 22432). Dagstuhl Reports. 2022;12(10):175-206. Available from: https://doi.org/10.4230/DagRep.12.10.175.

[2] Stab C, Kirschner C, Eckle-Kohler J, Gurevych I. Argumentation Mining in Persuasive Essays and Scientific Articles from the Discourse Structure Perspective. In: Cabrio E, Villata S, Wyner AZ, editors. ArgNLP, Forlì-Cesena, Italy, July 21-25, 2014. vol. 1341 of CEUR Workshop Proceedings. CEUR-WS.org; 2014. Available from: https://ceur-ws.org/Vol-1341/paper5.pdf.

[3] Xian X, Soon S, Feng T, Lewis J, Fong N. A powell optimization approach for example-based skinning in a production animation environment. In: Computer Animation and Social Agents. vol. 12; 2006. p. 2006.

[4] Lauscher A, Glavaš G, Ponzetto SP. An Argument-Annotated Corpus of Scientific Publications. In: Slonim N, Aharonov R, editors. Proceedings of the 5th Workshop on Argument Mining. Brussels, Belgium: Association for Computational Linguistics; 2018. p. 40-6. Available from: https://aclanthology.org/W18-5206.