Guided Linguistic Annotation of Argumentation through Visual Analytics

Fabian SPERRLE [b], Mark-Matthias ZYMLA [a], Mennatallah EL-ASSADY [b], Miriam BUTT [a], Daniel KEIM [b]
[a] *Department of Linguistics, University of Konstanz*
[b] *Department of Computer Science, University of Konstanz*

Abstract. We present a mixed-initiative approach to interactive annotation of argumentation, a typically time-consuming manual task. Our system facilitates the process by suggesting which fragments of text to annotate next. Suggestions are sourced from pre-annotations and user-preferences that are learned over time. Unused suggestions decay over time, reducing the amount of necessary interactions, while providing additional training data to the system. We show the effectiveness of the system for argument annotation according to Inference Anchoring Theory. The duality of suggestion sources and novel approach to suggestion decay are broadly applicable in linguistic annotation.

Keywords. visual analytics, argumentation annotation, guidance, interface

1. Introduction

High-quality annotated data is a constant concern in the domain of computational linguistics. On the one hand, manual annotation is laborious and time-consuming. Annotation schemes are often complex, and annotation guidelines unclear. On the other hand, fully automated annotation often results in lower quality, e.g., due to sparse training data. Argumentation annotation, in particular, is complex and includes several subtasks like text fragment annotation, labeling, and relationship extraction. Annotators can frequently switch between tasks and need to keep track of what they annotated so far. Further, they analyze the annotated arguments to identify patterns. The number of involved subtasks makes argumentation annotation a prime target for the integration of visual analytics techniques. Visual analytics is a research field that aims to include human judgment in analysis processes through visual representation and various interaction techniques [1].

In this paper, we present a general visual analytics approach to the guided acquisition of high-quality linguistic annotations of arguments. To bridge the gap between manual and automated annotation, we support users with adaptive suggestions of what to annotate next. The system learns user preferences over time to improve the quality and relevance of suggestions. Additionally, we provide distinct interface layers for the different annotation tasks that we connect with visually smooth, semantic transitions. The different views help to avoid information overload caused by large argument graphs. We exemplify the approach in *VIANA*, a mixed-initiative system for argumentation annotation according to Inference Anchoring Theory [2] (IAT). The system provides non-disruptive annotation sug-

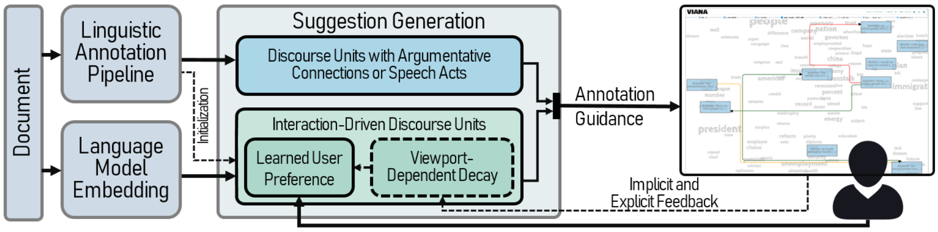


Figure 1. The annotation workflow in VIANA: The system generates suggestions based on linguistic pre-annotation and learns from user interaction over time. Users can provide explicit feedback by accepting or rejecting suggestions. Additionally, ignored suggestions decay over time, providing implicit relevance feedback.

gestions on two different tracks: a rule-based linguistic pre-annotation [3] and a weighted similarity model based on BERT [4] embeddings. Figure 1 shows the annotation workflow utilizing both types of suggestions. More generally, the system offers functionality for segmenting and labeling primary data, as well as introducing relational information between annotation objects—core components of many linguistic annotation schemes [5].

This paper contributes a showcase of our visual analytics technique for guided annotation on multiple (visual) layers [6]. Further, it highlights the applicability of user guidance to argumentation annotation and related text annotation tasks.

2. Background & Related Work: Argumentation Annotation & Interfaces

In the domain of argumentation annotation, the most prevalent system is OVA+ [7, 8], supporting large-scale argument map collection via AIFdb [9]. OVA+ is compatible with various variants of the AIF and, in particular, supports annotation according to IAT that we use to illustrate our system. However, it is a purely manual annotation system and does not provide automated support for annotators. BRAT [10] is a more general annotation system that features a wide array of capabilities for linguistic annotation. It can be extended with packages such as PAL [11] to add support for active learning components. Both OVA+ and BRAT provide a graphical annotation interface and have been informed by tools like GATE [12]. However, their efforts are not comparable to VIANA's visual analytics capabilities that enable the guided annotation in different task-specific views and layers. *Annotate* [13] provides an early version of a semi-automatic system for "syntactic annotation of natural language corpora" that frequently elicits explicit user feedback and integrates it into the automatic analysis process. In contrast, VIANA does not interrupt the user in their annotation and relies on implicit user feedback instead. Furthermore, it uses a novel approach to viewport-dependent suggestion decay to ensure the relevance of the considered implicit feedback.

IAT is used to annotate dialogue data and, specifically, debates [14]. The main focus lies in describing how argumentation is *anchored* in dialogue. For this purpose, IAT differentiates between the structure of the dialog (locutions and transitions between them) and the structure of argumentation (propositions connected via argumentative relations such as *inference*, *conflict* or *rephrase*). The dialogue- and argument-structure are linked via illocutionary connectors, which encode the illocutionary forces in the dialogue [2].

The annotation of argumentation in accordance with IAT is characterized by five high-level tasks, as identified by Sperrle et al. [6]. Annotators begin by closely reading the text at hand (T1) to gain a general understanding of its content. This phase may also

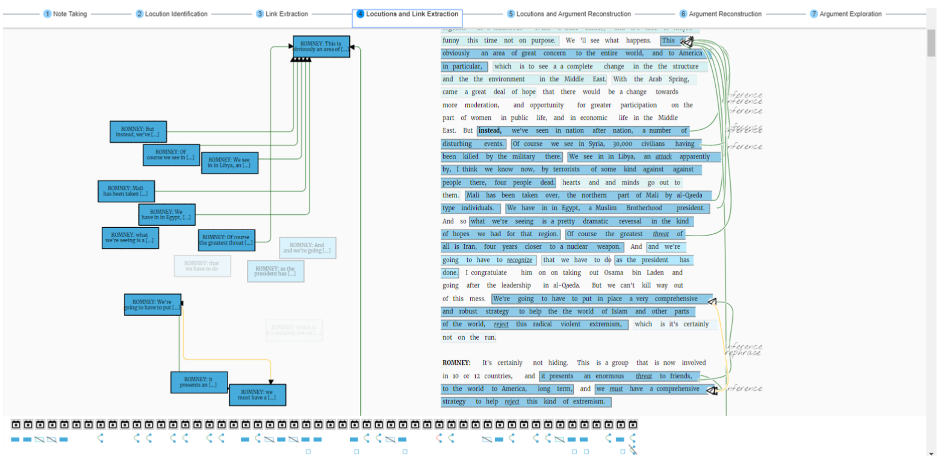


Figure 2. Overview of the VIANA system.

include note-taking to structure the following annotation process. The second task is the identification of *locutions*, fragments of text that contain claims which make up arguments (T2). The argument structure is encoded in terms of links between propositions that are anchored in locutions (T3). One of the main benefits of IAT is the translation of dialogue into a *logical structure* by reconstructing locutions into corresponding propositions (T4). Reconstruction entails, among other things, the resolution of anaphora and the rewriting of rhetorical questions into assertions. This step creates elements with explicit propositional content, thus providing a level of normalization to the argument graph. Finally, the annotation process ends with an exploration of the result (T5), transitioning annotators to the analysis phase. While task T4 is specific to the annotation of argumentation in IAT, the four other tasks are common in various linguistic annotation tasks.

3. The VIANA System

Traditionally, IAT graphs are visualized with locutions on the right and propositions on the left-hand side. The illocutionary connectors are shown between locutions and propositions. As VIANA [6] offers independent views for all tasks that are not necessarily on screen at the same time, illocutionary connectors are represented as badges and shown next to their respective left-hand side elements. In the system, all views are connected with visually smooth, semantic transitions to support switching views without losing context.

Note Taking The note-taking view enables slow analytics [15] and represents the "distraction-free" mode of the system. Users can read the text and add notes to interesting fragments, understanding the data before beginning the IAT annotation.

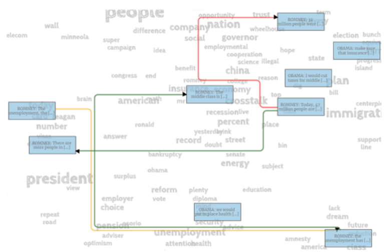
Locution Identification The locution identification view enables the annotation of locutions directly in the text view: users click and drag to highlight and annotate the text (dark blue highlights). VIANA displays suggestions of what to annotate next (see section 4) to structure the annotation process and support annotators. Suggestions are sourced from linguistic pre-annotation (light blue) or based on learned user preferences (teal). For the latter, the opacity encodes the system's confidence in the suggestion. To accept or reject suggestions, users click the respective buttons in a tooltip. Ignored suggestions decay over time to free users from constantly rejecting suggestions during the learning phase of the system.

The locution identification view can also be used to introduce relations between both the identified locutions and their associated propositions. Both transitions and propositional relations, such as inferences and conflicts, can be added by dragging and pressing the control or shift key, respectively. To add undercuts, new relations are drawn to the label of an existing edge, while linked arguments are constructed by drawing to the arrowhead of an existing relation. The type of added propositional relations (inference, conflict, or rephrase) can be cycled by double-clicking the relation or via a tooltip.

Relationship Extraction The relationship extraction view displays a graph of shortened propositions at the exact position of their corresponding locutions (where applicable). This supports users in transitioning through the layers, as no nodes have to be moved on the screen. This view is tailored towards the addition of further relations that have not yet been added when extracting the locutions. It provides the same interactions for adding relations that are also available in the locution identification view and were described in the previous section. However, as the underlying text is no longer visible, annotators can focus on the task at hand with fewer distractions.

Locution Reconstruction Reconstruction of locutions into propositions entails changing the content of extracted locutions to contain some explicit propositional content. More concretely, the goal is to provide a sentence that is fully understandable without any context. To support this task, the proposition reconstruction view expands the shortened graph nodes from the relationship extraction view and displays the full content of the propositions. Additionally, propositions are temporally untangled and displayed such that they can be read as a timeline from top to bottom. To reconstruct a proposition, users double-click on any of the graph nodes to edit the text in-place.

Argumentation Exploration The argument exploration map is constructed based on a semantic concept space [16]. A concept space identifies the key concepts of the underlying text and provides a spatialization thereof. We then gather the concepts associated with the identified propositions and position them in the provided spatialization. This view gives a quick overview of the debate's content and can be used to check the annotation progress. Furthermore, users expressed that they would use it as a "sanity check" to identify whether they might have missed core concepts in the annotation.



4. Annotation Guidance

To speed up the annotation of text and structure the process, VIANA provides user guidance in the form of suggestions for text fragments that may require annotation. This guidance is provided in two different "tracks". The first track suggests discourse units that have been identified by a linguistic pre-annotation pipeline [3] and are connected with relations of type *conclusion*, *reason*, *condition* and *consequence* or contain a speech act of type *agreement* or *disagreement*. The second track learns a weighted similarity between annotations over time and is relevant wherever no discourse markers have been identified during the pre-annotation. The suggestions of both tracks are visualized in blue and teal, respectively, to enable users to identify the provenance of each suggestion easily. Such transparency is important for judging the quality of suggestions and fosters trust calibration.

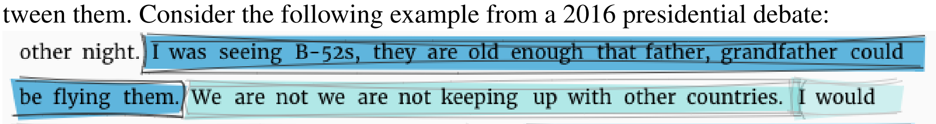
Implicit and Explicit Feedback Once the user begins to select or reject suggestions or provides manual annotations, the system learns from those explicit interactions by refining a weighted text similarity model. It then proceeds to suggest fragments that are similar to those that have already been annotated. Additionally, implicit user feedback is gathered through a novel approach to viewport-dependent suggestion decay. Whenever users ignore provided suggestions but interact closeby on the screen, the system reduces its confidence in the ignored suggestion. As text annotation typically requires knowledge of the context of a suggestion to judge its validity, the confidence loss decreases as the distance increases. Once the system confidence in a suggestion falls below a given threshold, the underlying text fragment is turned into a negative training sample for the guidance component, essentially mimicking manual rejection of the suggestion through the user. This concept enables gathering larger amounts of training data from little interaction.

Guidance Triggers and Adaptation To avoid distracting annotators from their task, all guidance in VIANA is provided without explicit triggers: users are continuously presented with suggestions without having to request them first, and all interactions influence future suggestions. This implicit guidance model enables the system to provide a high number of suggestions without interrupting the user's workflow to elicit feedback or serve more suggestions and is thus particularly suited to complex annotation tasks. In the background, VIANA relies on a co-adaptive guidance process [17]: the system suggests annotations, aiming to teach the user. At the same time, users provide feedback to those suggestions or create manual annotations, allowing the system to learn a user-specific guidance model. This user-specificity and personalization are important as different users require different guidance, for example, due to different levels of expertise. A previous evaluation of the system revealed that users adapt their guidance track selection strategy over the course of an annotation session [6], placing more and more trust in the suggestions based on learned similarity as the system adapts to their annotation style over time.

5. Application Examples

As we have previously evaluated VIANA in a series of expert user studies [6], we now focus on describing some annotation instances pertaining to the different tasks.

Local Arguments We illustrate the annotation of arguments of the form "*premise* supports *conclusion*" with two claims (premise and conclusion) and a relation (support) between them. Consider the following example from a 2016 presidential debate:



Only the first claim, the premise of the argument, has been annotated so far (indicated by the deep blue color). The conclusion is then automatically suggested (teal color) as an annotation based on its similarity to previously annotated locutions. After accepting the suggestion, the user can finish the annotation by linking the two claims with a support relation. This example showcases both the suggestion functionality, as well as the ease of annotating simple arguments directly in the text view.

Long-Distance Arguments Although not as common, long-distance connections between claims occur regularly in IAT annotations. Figure 2 and Figure 3 both show instances of long-distance argumentation. Such relations are easy to overlook in the text window. However, the *link extraction* view (see Figure 2, left-hand side), as well as the *argument exploration* view (Figure 1, right-hand side) counteract this: *dangling* locutions that are not yet connected with relations can be identified in these graph-based views.

Taskwise Argumentation Annotation Figure 3 shows an advanced stage of annotation at which most locutions have already been embedded in arguments. Next, the locutions need to be reconstructed in accordance with the IAT specification (T4). The *Argument Reconstruction* view allows the annotator to edit a proposition in a focused view, containing only other propositions as context. Removing additional information, like locutions or the underlying text, can facilitate the task as it reduces the risk of information overload. Whenever more context is necessary, a layer combining the *Argument Reconstruction* and *Location Identification* views directly relates propositions to their associated locutions in the original text. This facilitates, in particular, the process of anaphora resolution, which is one of the more challenging issues of locution reconstruction.

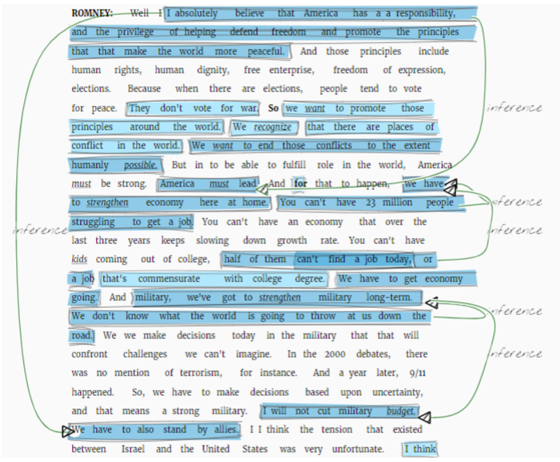


Figure 3.Annotation in the Locution View.

6. Conclusion

We have presented VIANA, a visual analytics approach to task-dependent text annotation. Our approach is designed to ensure adherence to structured annotation guidelines. To focus the interactions, tasks are separated into different views [6], reducing the risk of information overload. This separation of views also enables tackling multi-level annotations in different task orders, making the system suitable for various annotation styles. Furthermore, the system adapts to and learns from user input. It provides guiding suggestions to potentially speed-up the annotation process and increase its consistency. Such a mixed-initiative approach attempts to learn the users' rationale and annotation model, bringing us a step closer to the automation of such annotation processes.

The presented concepts are expandable beyond the scope of annotation. They showcase the benefits that the inclusion of visual analytics can bring to tasks like argumentation analysis that require the adherence to structured guidelines, as well as the flexibility of human feedback and input to succeed.

References

[1] Keim DA, Mansmann F, Schneidewind J, Thomas J, Ziegler H. Visual Analytics: Scope and Challenges. In: Simoff SJ, Böhlen MH, Mazeika A, editors. Visual Data Mining: Theory, Techniques and Tools for Visual Analytics. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer; 2008. p. 76–90.

[2] Budzynska K, Reed C. Whence Inference. University of Dundee Technical Report. 2011.

[3] Hautli A, El-Assady M. Rhetorical Strategies in German Argumentative Dialogs. Argument & Computation. 2017 8;(2):153–174.

[4] Devlin J, Chang MW, Lee K, Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding; 2018.

[5] Ide N, Pustejovsky J. Handbook of Linguistic Annotation. Springer Netherlands; 2017.

[6] Sperrle F, Sevastjanova R, Kehlbeck R, El-Assady M. VIANA: Visual Interactive Annotation of Argumentation. In: Proceedings of IEEE Conference on Visual Analytics Science and Technology (VAST); 2019. p. 1–12.

[7] Reed C, Janier M, Lawrence J. OVA+: An Argument Analysis Interface. In: Computational Models of Argument: Proceedings of COMMA. vol. 266; 2014. p. 463.

[8] Lawrence J, Visser J, Reed C. An Online Annotation Assistant for Argument Schemes. In: Proceedings of the 13th Linguistic Annotation Workshop. Association for Computational Linguistics; 2019. p. 100–107.

[9] Lawrence J, Bex F, Reed C, Snaith M. AIFdb: Infrastructure for the Argument Web. In: Computational Models of Argument: Proceedings of COMMA; 2012. p. 515–516.

[10] Stenetorp P, Pyysalo S, Topić G, Ohta T, Ananiadou S, Tsujii J. BRAT: a Web-based Tool for NLP-assisted Text Annotation. In: Proceedings of the Demonstrations at the 13th Conference of the European Chapter of the Association for Computational Linguistics. Association for Computational Linguistics; 2012. p. 102–107.

[11] Skeppstedt M, Paradis C, Kerren A. Pal, a Tool for Pre-annotation and Active Learning. Journal for Language Technology and Computational Linguistics. 2017;31(1):91–110.

[12] Cunningham H, Maynard D, Bontcheva K, Tablan V. GATE: an Architecture for Development of Robust HLT applications. In: Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics. Philadelphia, Pennsylvania, USA: Association for Computational Linguistics; 2002. p. 168–175.

[13] Plaehn O, Brants T. Annotate – an Efficient Interactive Annotation Tool. In: Proceedings of the 6th Conference on Applied Natural Language Processing; 2000. .

[14] Budzynska K, Janier M, Reed C, Saint-Dizier P, Stede M, Yaskorska O. A Model for Processing Illocutionary Structures and Argumentation in Debates. In: LREC; 2014. p. 917–924.

[15] Bradley AJ, Sawal V, Collins C. Approaching Humanities Questions Using Slow Visual Search Interfaces. In: Proceedings of the VIS4DH Workshop at IEEE VIS 2019; 2019. p. 1–4.

[16] El-Assady M, Kehlbeck R, Collins C, Keim DA, Deussen O. Semantic Concept Spaces : Guided Topic Model Refinement using Word-Embedding Projections. IEEE Transactions on Visualization and Computer Graphics. 2020 Jan;26(1):1001–1011.

[17] Sperrle F, Jeitler A, Bernard J, Keim DA, El-Assady M. Learning and Teaching in Co-Adaptive Guidance for Mixed-Initiative Visual Analytics. In: Vrotsou K, Turkay C, editors. EuroVis Workshop on Visual Analytics (EuroVA). The Eurographics Association; 2020. .